University of Warsaw
Interdisciplinary Doctoral School

SCALE-FREE COMMUNICATION?

An investigation of the use of the concept "communication" in biology and cognitive sciences

WIKTOR ROROT

Doctoral dissertation prepared under the supervision of
PROF. DR HAB. JOANNA RĄCZASZEK-LEONARDI
Faculty of Psychology, University of Warsaw

and

DR HAB. MARCIN MIŁKOWSKI, PROF. IPS PAS
Institute of Philosophy and Sociology, Polish Academy of Sciences

Warsaw, December 2025

COLOPHON

This document was typeset using LaTeX with the `classicthesis` package developed by André Miede. The bibliography was processed with Biblatex.

The cover image is by Jacques Fabien Gautier d'Agoty, illustrating the *Anatomie de la tête* by Duverney, published in 1748. Available in the public domain through Wellcome Collection.

Text and headings use Sebastian Kosch's Crimson Roman. Monospaced text uses Jim Lyles's `Bitstream Vera Sans Mono` ("BeraMono"). Formulas are typeset with Euler font, designed by Hermann Zapf and adapted to OpenType by Khaled Hosny and Daniel Flipo.

Scale-Free Communication? An investigation of the use of the concept "communication" in biology and cognitive sciences

SUMMARY

The goal of the thesis is to address the increasing reliance on communication-centred language across life and mind sciences.

A growing number of researchers in different fields of biology and cognitive science employ the term “communication” in their descriptive and explanatory practices. Processes that are subsumed under this notion include—but are not limited to—interactions between single cells, cell ensembles, organisms of varying complexity, both unicellular and multicellular, and at times between animals and artifacts (such as computers).

However, the complex roles the term “communication” plays in the scientific practice of biologists has hitherto received limited attention in philosophy. At the same time, a growing number of researchers view the conceptual framework of communication as fruitful. This leads to proposals (such as the “scale-free biology” approach proposed by Michael Levin and Christopher Fields) that seek to integrate insights from different branches of biology through the terms of “information” and “communication”, expanding beyond the synthesis provided by previous approaches. For this reason, it is increasingly important to scrutinize what “communication” means across the different contexts where it is used, what role it fills in the scientific practices that rely on the term, and what implications its use has for the way we understand biological phenomena.

To answer these questions, I employ the tools of digital philosophy of science to examine a dataset of over 1 million open-access scientific articles from the relevant areas of biology and cognitive science, sourced from the Semantic Scholar Open Research Corpus. The adopted multi-level methodology proceeds in three main steps: first, references to “communication” in the dataset are identified through semantic search. Second, these passages are investigated using topic modelling and collocation analysis. Finally, the results of this computational analysis are analysed in parallel with a close reading of textual data to formulate a defintion of the concept and a detailed account of the role it plays in conceptual practice of science.

Based on the analysis of the data, I submit that there is a continuity in the use of communication notions as applied to phenomena across spatiotemporal scales of life, i.e., at the scale of individual cells, tissues, and whole organisms. I support this proposal with an argument by negation. Using the notion of “patchwork concepts” I show that “communication” is in fact not polysemous and has a core meaning that is shared by the different uses of the concept. I thus propose a novel definition of biological communication. This shared meaning of “communication” allows the concept to play the same role across scales, by describing the causal structure of biological processes, which shares the same key components: sender, receiver, and signal.

To highlight the epistemic benefits of my proposal, I explore how “communication” relates to formal frameworks of information theory, highlighting some of the idealizations these formalisms introduce. Ultimately, I argue for a realist reading of “biological information”, as compatible with the proposed account of “communication”.

Further, I analyze how scientists ascribe meanings to the signals at various scales. Interestingly, despite a shared understanding of “communication”, the practices of identifying semantics of signalling differ across scales. I show the distinctions between the meaning at organism- and cell-level, drawing on existing accounts from philosophy of mind and philosophy of biology. Based on analogies between these models, I extend the concept of “meaning as control” originally proposed within ecological psychology, to capture the connections between the scales.

The dissertation develops a novel account of biological communication covering the whole scope of the use of this notion in biological and cognitive science research. The proposed account of “scale-free communication” can inform investigations into the biological underpinnings of cognitive abilities, including within the emerging field of basal cognition. As communication functions both as a conceptualization of biological phenomena, and a set of focused formal models, the study of this notion offers important insights into how theoretical reasoning intersects with operationalization and formalization in actual scientific practice. Lastly, the methodology used can be extended to other cases, allowing for the study of conceptual frameworks of science at a large scale and across disciplinary boundaries.

Keywords

communication, information, animal signalling, cell signalling, digital philosophy of science, basal cognition, conceptual framework, philosophy of science in practice, naturalistic philosophy of science, digital humanities

ACKNOWLEDGMENTS

As I'm putting final touches to the dissertation, and thinking back on the route I have taken to get to this point, I cannot but wonder about the spur-of-the-moment, serendipitous choices I have made at various stages of my life that have eventually led me to where I am now. This would never have been possible if not for the people I have met on this journey.

I would like to thank my former mentors: Grażyna Strużyna-Handke, Dagmara Burnus-Szymczyk, Maciej Maryl, Joanna Komorowska-Mach, and Paweł Gładziejewski, for all the support and belief you've had in me, and for the various—sometimes random—paths I have taken under your guide, which eventually fit together surprisingly well.

Crucially, my supervisors, Joanna Rączaszek-Leonardi and Marcin Miłkowski, have given me the freedom and tools to explore my interests, while at the same time keeping me (somewhat) on track. They have always shown patience with the various applications and side projects, and continuously fostered a unique intellectual environment which made interdisciplinary thesis like this one possible. I couldn't have hoped for better PhD supervisors.

During my studies, I have been able to meet a number of wonderful people, to discuss their work with them, to collaborate, or simply enjoy our time together. This list is too long to include here—but I want to thank you for making this journey so pleasurable. Of the various conferences, summer schools, study visits I have participated in, this dissertation has beeen most directly influenced by:

- Diverse Intelligences Summer Institute which I have attended in 2019 in St. Andrews and in 2021 online—thanks to DISI leadership: Erica Cartmill, Jacob Foster, Kensy Cooperrider, and my fellow Fellows: Taylor Beck, Daniel Czegel, Ben Falandays, Hamza Giaffar, Judit Mokos, Anastasia Nikoulina;
- the online study visit to Michael Levin's Lab at Tufts University in 2022—with thanks to Michael Levin for this wonderful opportunity;
- the visit to Charles Pence's group at the Université catholique de Louvain in 2023—with thanks to Charles Pence for his time and help;
- and the Writing-Up Fellowship at the Konrad Lorenz Institute in Klosterneuburg in 2024-2025—thanks to KLI leadership: Isabella Sarto-Jackson, Barbara Fischer, Philipp Mitteroecker, Gerd Mueller, and my cohort: Olesya Bondarenko, Joyshree Chanam, Elis Jones, Merin Joji, Onerva Kiianlinna, Paul Knabl, Marina Knickel, Nina Kraus, Jules Macome, Somya Mani, Sean Pears, Enrico Petracca, Aleksa Ratarac, Ludo Schoenmakers, Hari Sridhar, Eva Zaffarini.

Among the many friends I've made in academia, I want to particularly thank David Harrison and Urte Laukaityte, my lovely collaborators, as well as Marco

Casali, Alejandro Fábregas-Tejeda, Marcin Furtak, Mateusz Kostecki, Nedah Nemati, Aleksandra Piejka, Guido I. Prieto, Martina Radice, Natalia Rutkowska, Lorenzo Sartori, Krzysztof Sękowski, Nat Valadzko, Justyna Wiśniewska, Julian Zubek.

I have enjoyed endless support and patience of my friends, which helped me through the most difficult parts of these years: Adrianna Biernacka-Gemel, Łucja Doradzińska, Paweł Gorczyński, Oskar Karski, Jakub Musiał, Dominika Pękalska, Przemysław Pożar, Aleksandra Sołtysik, Ewelina Tołwińska, Paulina Woźniczka, and many, many more. Thank you for your love and kindness.

Finally, I want to thank my family for their unconditional support and for allowing me to become who I am now, especially my parents and sisters for the caring home I grew up in and their unshakeable belief in me, even as I have decided to overlook the more reasonable career options.

Among the many happy decisions I have made, one that I particularly enjoy the outcomes of on a daily basis is the choice to share my life with my partner-turned-wife, Kasia. Thank you for our life.

Let it be known that the cats, Lesław and Masław, have NOT contributed to the dissertation, but they will be thanked nonethelessdauwihadl."',qqqqqqqqqs/d.f,.

FUNDING STATEMENT

The completion of this dissertation was possible thanks to the generous funding of:

- National Science Center (Poland) as part of Opus grant "Developmentally informed agent-based modeling of symbolic constrains in interaction" (2018/29/B/HS1/00884) awarded to Joanna Rączaszek-Leonardi, where I was a PhD student researcher (October 2020-January 2023);
- National Science Center (Poland) as part of Preludium grant "Investigation of the use of the concept "communication" in biology and cognitive sciences" (2022/45/N/HS1/02434; supervisor: Marcin Miłkowski), where I am a principal investigator (from February 2023, planned conclusion: July 2026);
- University of Warsaw (Poland) through the doctoral scholarship awarded as part of the Excellence Initiative—Research University in 2023;
- Foundation for Polish Science through the START Scholarship awarded in May 2024;
- Ministry of Science and Higher Education (Poland) through the Scholarship for outstanding young researchers, awarded in June 2024 for the period of October 2024-September 2027;
- Konrad Lorenz Institute for Evolution and Cognition Research (Austria) through the Writing-Up Fellowship which enabled my stay at the KLI in the period of October 2024-March 2025.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF BOXES

1

INTRODUCTION

The application of the language of the newly-born information theory to biological and psychological research in the 1950s and 1960s has been one of the most significant paradigm shifts in contemporary life sciences (J. H. E. Cartwright et al. 2024; Gatenby and Frieden 2007; Gleick 2011; Kull 2000). Its impact on the empirical and epistemic practices of these disciplines can only be likened to the pervasive molecularization of biological studies (for molecularization, see Chadarevian and Kamminga 2003; Morange 2020). Indeed, the concept of INFORMATION has been extensively discussed by philosophers, across the three major contexts of biological and cognitive research it appears in: in genetics, cognitive science (including psychology and neuroscience), and ecology and animal behaviour. Yet, these extensive theoretical accounts have overlooked that along with “information”, biology has picked up the closely related terms of “signalling” and “communication” that soon assumed a life of their own. COMMUNICATION (independently of information) is now routinely invoked in life and cognitive sciences to describe processes as different as: the coordination in bacterial biofilms (quorum sensing), cell differentiation during metazoan development (bioelectric communication), operation of the nervous system—and its relation to other types of tissues within the organism (neural signalling, gut-brain communication), a host of different social behaviours of animals—from alarm calls, to the complex games of authority and power in human politics.

This intricate landscape of uses prompts the question: what is communication? Do bacteria and humans really engage in the same process, or is it an (unfortunate) metaphor or analogy?

To much surprise, there is no attempt to answer these questions in the literature. Indeed, there’s a crucial asymmetry in the discussions of BIOLOGICAL COMMUNICATION. Almost all the theoretical appraisals of the concept, both in philosophy of biology and within biology itself, focus only on one end of the spectrum: they deal with communication between animals, sometimes extending to bacteria or plants, but remaining at the scale of communication between *organisms*. Biologists, however, commonly ascribe COMMUNICATION labels to processes and mechanisms at the cellular or even subcellular scale. How are the gaps between this variety of spatial scales, from the very small to the reasonably large, bridged in practice, what motivates the use of the same notion in such different contexts?

The question of the parallels in the uses of “communication” in life and cognitive sciences across the scales of biological organization is the central research problem of this thesis. Indeed, beginning the research project, I have hypothesized that it will be possible to indicate that there is a single concept of BIOLOGICAL communication underlying these disparate uses in the literature. To operationalize this hypothesis, I propose the notion of "scale-free communication", motivated by the concept of fractal similarity or self-similarity. What this term encapsulates is the suggestion that communication captures certain organization of biological processes, which can be realized at various scales of living systems.

This terminological choice is motivated by the recent use of the notion of "scale-free" by Michael Levin. Levin attempts to provide an account of cognition as a scale-free process (M. Levin 2019). According to this proposal, the complex cognitive capacities of systems equipped with a central nervous system can be (explanatorily) decomposed into a plethora of simpler cognitive systems implemented at lower scales of biological organization, down to individual cells (see also M. Levin and Dennett 2020; Chis-Ciure and M. Levin 2025). This perspective has motivated a broader suggestion, where Levin with Chris Fields argue for a novel synthesis of various areas of biological research. They propose to go beyond the evo-devo (e.g., Gould 1977; Oyama 2000) and the extended evolutionary synthesis (e.g., Pigliucci and Müller 2010b, 2010a) through "the language of communication, inference, and information processing" (Fields and M. Levin 2020, p. 1). Indeed, some authors argue for yet greater significance of communication. J. Mark Bishop and Slawomir Nasuto attempt to replace the computational view in neuroscience with a communicational alternative (e.g., Bishop and Nasuto 1999; Nasuto, Dautenhahn, and Bishop 1999), together with alternative neural network architectures highlighting the potential of communication. More broadly, Günther Witzany (2019; 2020) has developed a research programme which seeks to establish communication as a central feature organizing life, its main characteristic. Finally, the field of biosemiotics (e.g., Barbieri 2008; Emmeche and Hoffmeyer 1991; Hoffmeyer 2008), considers the central importance of various forms of semiotic—sign-based—processes across biology, including communication.

While these speculative considerations partially motivate my interest in biological communication, they are somewhat marginal for the main argument advanced here and will be bracketed for the majority of the discussion in the thesis. Nevertheless, if the hypothesis of "scale-free communication" is correct, this is where the largest epistemic pay-off of the current project may lie. For this reason, I eventually come back to the theoretical importance of the developed understanding of communication and consider its implications. In a way, a part of what this thesis accomplishes is a "de novo" recovery of the biosemiotic recognition of the role of communicative processses directly from the conceptual practice of contemporary life and mind sciences.

Importantly, it is the expansion of the scope of "communication" (and relatedly, "information") that distinguishes biology and cognitive sciences from other domains that use these concepts. Linguistics, communication studies, sociology, etc. apply the notion within its dictionary meaning, to exchanges between humans. As such, the analyses conducted here focus on this special case of biological and cognitive-scientific use of the concept.

As it will turn out in the following chapters, empirical data from the (elements of) scientific practice seem to agree with the hypothesis of "scale-free communication", though—as might be expected—the data introduce some further complexities into the understanding of the connections between scales. Indeed, while the main question of the thesis has speculative implications, the immediate goal is much more "down to earth", in that I aim to recover the meanings of "communication" and related terms, as well as their descriptive and explanatory roles in the contemporary research across the life and mind sciences, directly from the textual scientific practice. To this end, I employ computational methods of digital philosophy of science to study a corpus of over 1~million scientific articles, connecting those tools to more standard, interpretative philosophical methodologies.

Digital philosophy of science is a novel approach within the naturalistic philosophy of science. It seeks to overcome some limitations of the other approaches of that field, in particular the arbitrariness inherent in the method of case studies. To this end, digital philosophy of science harnesses the tools developed within the digital humanities (Berry 2012a) and applies them to study the corpora of scientific texts, with the goal of answering some standard questions of philosophy of science. The first building blocks for this approach have been laid down by Paul Thagard with his proposal of "computational philosophy of science" (Thagard 1988). Thagard's approach viewed scientific processes as computational and, crucially, argued for a possibility of a link from a descriptive to normative conclusions in a philosophical enterprise.

However, the proper history of digital philosophy of science, understood as the use of methods from computational linguistics and natural language processing to study digital (or digitized) outputs of scientific processes (especially scientific articles), has begun a little over a decade ago. One of the early attempts was the study conducted by James Overton on a year-worth of articles published in *Science*, where he sought to establish the importance of "explanation" as an aim of science (Overton 2013). Other projects soon followed, expanding the methods to cover the range of tools used within digital humanities more broadly—a slightly older, already established approach within history, linguistics, and literature studies (among others), which explores textual sources with computational tools to answer a plethora of disciplinary questions that previously remained beyond the epistemic reach of researchers, or required laborious manual procedures (e.g., Berry 2012b; Tilton, Mimno, and J. M. Johnson 2024).

1.1 OUTLINE OF THE THESIS

In the remainder of the introduction, I provide several important preliminary points regarding BIOLOGICAL COMMUNICATION, including a standard definition, and situate the discussions presented in the thesis in the context of the more (philosophically) established debate surrounding BIOLOGICAL INFORMATION, as there are several important intersections between these areas.

Next, in chapter 2, I discuss the methodological context of the project. I position it within the field of naturalistic philosophy of science, adjacent to the approach of philosophy of science in practice, though there are significant limitations of how much scientific practice remains visible in scientific publications—a topic I cover in depth in section 2.1. Digital philosophy of science, as I view it, emerges as a particular set of tools. Inspired by the developments in computational linguistics and digital humanities more broadly, it allows philosophy of science to overcome some of its restrictions and expand its reach. Nevertheless, it comes with its own set of challenges. In section 2.2 and 2.3, I offer a broad review of current state-of-the-art work within the digital philosophy, introducing its various tools, and explore its constraints. I conclude the chapter, and the introductory part of the thesis, in section 2.4 and 2.5 by providing a detailed exposition of the methodology of the studies reported in this thesis.

Chapters 3-5 constitute the main part of the dissertation. In these chapters I report and interpret the results of the computational corpus studies that explored the use of the terms related to BIOLOGICAL COMMUNICATION. Chapter 3 seeks to establish how researchers across life and cognitive sciences understand communication. To this end I develop an argument by contradiction. I assume that communication is a polysemous notion, and that it captures distinct, even if related, processes across scales. I operationalise this drawing on the notion of “patchwork concepts” (M. Wilson 2006; Haueis 2024), and analyse the use of “communication” in the corpus through this lens. While indeed there are distinct patches—identified in terms of domains, scales, techniques, and properties the techniques target—the textual results highlight that the patches are not just locally connected. The analysis highlights that the general definition-schema introduced below (section 1.2) is not just a “general reasoning strategy”, providing a “pragmatic unity to the different uses of a patchwork concept” (Haueis 2024, p. 744). Instead, these different uses hinge on a shared understanding of the concept of communication—differently from what the perspective of “patchwork concepts” assumes. I explore this core meaning of the concept, and conclude that indeed BIOLOGICAL COMMUNICATION is a scale-free notion, applicable at all scales of biological organization. Nevertheless, there are some important limitations to the account of COMMUNICATION recovered directly from the scientific literature, hence I consider the definition proposed in the light of a series of case studies drawn from the literature, to further finesse the understanding of what BIOLOGICAL COMMUNICATION is, and how it relates to other causal phenomena.

In chapter 4, I explore specifically how this view of communication relates to information theory—and how information theory features in the biological discussions of communication across the scales. This analysis allows me to highlight several misconceptions in extant philosophical treatments of these questions. More importantly, however, I focus on numerous central assumptions that formal frameworks of communication make, and which are not met in the biological realm. Considering these boundaries for the formalisms in detail, I argue nonetheless for a realist interpretation of the "information-talk" in the context of BIOLOGICAL COMMUNICATION. I propose a perspectival realist view on the matter, which allows for the appreviation of epistemic benefits of the use of information-theoretic tools—despite the distortions these models introduce.

Finally, in chapter 5, I touch upon one of the central questions at the intersection of philosophy of biology, of cognitive science, and of mind: the naturalistic explanation of the origins of semantics. Motivated by the realization that at least apparently, meanings are imputed to signals across the scales of life, I use a number of leading naturalistic accounts of intentionality from the philosophical literature to uncover whether and how researchers in life and cognitive science ascribe meaning to biological signals. The results indicate that such "semantic practices" are widespread across the different domains covered by the dataset, but there are important differences in how such ascriptions are made with regard to the cell-level and organism-level processes. I analyse these different contexts through the lens of existing philosophical accounts of meaning to make sense of the scientists' practices, and identify the sources of this split. Ultimately, though, I arrive at a view which allows for a shared origin of the two types of semantic practices and casts their differences as related to the increasing complexity of the processes at the organism level, and the different research questions asked in the two contexts, rather than a more substantial, "metaphysical" distinction. This view is possible if we adopt the emerging idea of constraint-based semantics (e.g., Rączaszek-Leonardi 2012), which allows us to bridge the gap between the two different contexts visible in the dataset.

Chapter 6 summarizes the findings presented in this thesis to establish the role of the concept of COMMUNICATION for contemporary research in cognitive and life sciences. Having done this, I also consider the fine balance between a descriptive and a normative position that I've been trying to strike in this thesis, expanding on how it is best to understand the results I have achieved here. Chapter 6 is also where I finally allow myself to explore the very tentative and highly speculative connection of the current research project to the emerging programme of basal cognition (e.g., Lyon et al. 2021). I briefly consider some criticisms raised against that approach, and highlight how the results of the current research open up a path for basal cognition research to more directly tie into other areas of the study of cognition. Despite the necessarily sketchy form of those suggestions, I hope that they succeed at highlighting the epistemic impacts I envisage for this thesis.

1.2 VIEWS ON BIOLOGICAL COMMUNICATION

As I have already mentioned, philosophical reflection on the use of the term "communication" in biology has almost exclusively focused on the research on animal communication, as a problem area within the animal behaviour studies. Some researchers acknowledge that the term features also in biological research on smaller spatiotemporal scales of biological life. Nevertheless, there is no systematic treatment in that context, nor any clear consideration on how the larger-scale notion relates to these smaller scales—a research gap this project aims to fill.

In the context of human language, Ulrich Stegmann (2016) identifies in the existing literature four separate approaches to understanding communication:[1]

- code model,
- inference model,
- extended sense model,
- signalling model.

All these competing models emerged in philosophy of language and attempt to capture several important factors of language systems that become most evident in communicative interactions. Stegmann's classification—and indeed, a majority of work in this area—focuses on understanding what the success of linguistic communication consists in. This success quite naturally can be accounted for in terms of "understanding" (though this is often not the default philosophical frame used in that context, see Dahl, forthcoming). One of the earliest proposals, termed the "code model", views signals (or their components) as being associated with (mapped to) some propositions (in the logical sense) or thoughts (on an internalist interpretation). According to this approach, communication consists in the sender encoding the referent in an appropriate signal, and the receiver decoding the signal, according to the shared code. This model is quite simplistic and has important limitations—not least, the problems with indexicals, i.e., expressions which cannot be consistently mapped to some single referent, and are context-dependent (e.g., pronouns). Further, problems arise when we consider pragmatic components of communication, for example an ironic tone, which effectively transforms the conveyed sense.

These problems have been most famously addressed in the work of H. Paul Grice (1957; 1969; 1995), who is seen as one of the originators of the alternative "inference model." According to this model, to understand an utterance, the hearer attempts to infer the speaker's intention to convey a thought through that utterance. In this context, an individual expression acts as a piece of evidence the hearer can use to arrive at the speaker's intended meaning. This approach has been highly influential in providing a philosophically interesting framing of

---

[1] I follow Stegmann's treatment in this characterization, which is intended as a broad sketch to span the conceptual dimensions of the debates recurring throughout the thesis, and as such it is necessarily incomplete.

pragmatic components of communication, and motivated several further models of linguistic interactions, including the relevance theory (Sperber and D. Wilson 1995; Heintz and Scott-Phillips 2023; Scott-Phillips and Heintz 2023, this is covered in greater length in chapter 5). Nevertheless, this approach faces some important limitations, including the relatively high cognitive requirements with regard to the hearer's ability to get at speaker's meaning. More importantly, however, neither the code nor the inference model are able to fully account for the role of communication.

This last realization motivated the proponents of what Mitchell Green (2007) called the "extended sense model." The central observation introduced by this account is that the internalist orientation of code and inference models overlooks the importance of communication in providing hearer's access to the states of the world they would not be able to perceive otherwise. In that context the extended sense model supplements the former accounts by focusing on the relation of utterances to external states of the world that the hearer can epistemically access through that utterance. Nevertheless, this model also assumes relatively high-level cognitive abilities on the part of actors involved.

The final model Stegmann distinguishes—one that is primarily developed to bridge the gap between the animal and human communication systems—is the "signalling model" (e.g., Bennett 1979; Green 2007). This model views animal communication as a "more basic" form of human linguistic interactions (which are still, at least according to Bennett (1979), to be explained in broadly Gricean terms) which consists of information transmission: communication in that model is those behaviours which have a *function* of transmitting information (the notion of "function" is indeed an important concept in this context and I come back to it in chapter 5).

Beyond the positions noted by Stegmann, the "signalling model" is not the only attempt to explain communication as a set of behaviours shared by humans and non-humans. The philosophical account which has seen the largest uptake, both in philosophical discussion and directly in empirical research, has been offered by Thom Scott-Phillips (2008). Scott-Phillips begins by contrasting two opposed, in his view, approaches found in the literature. One tradition is the "signalling model" just discussed, which seeks to define animal communication in terms of information transfer. Another focuses on the adaptive benefits that (putatively) communicative interactions and behaviours confer upon the organisms involved—senders and receivers (this general distinction into information- and influence-based definitions of communication has come to be widely adopted, see e.g., Stegmann 2013; Frick, Bich, and A. Moreno 2019). Such an explicitly "informational" definition has been offered by Marc Hauser (2000, p. 6), who notes that "[s]ignals carry certain kinds of informational content, which can be manipulated by the sender and differentially acted upon by the perceiver."

The second view, in turn, is exemplified by one of the most widely used definitions of communication, offered initially by John Maynard Smith (2000), and improved upon in his later work with David Harper (Maynard Smith and

D. Harper 2003). These authors define communication through the notion of a "signal"—namely, the vehicle of a communicative interaction, but without the informational implications. For Maynard Smith and Harper (2003, p. 3), signal is "any act or structure which alters the behaviour of other organisms, which evolved because of that effect, and which is effective because the receiver's response has also evolved."

These definitions of communication were traditionally developed to capture several paradigmatic instances, including human communication, vervet monkeys' (*Chlorocebus pygerythrus*) alarm calls, or honey bees' waggle dance. They are usually benchmarked against a number of standard borderline cases, primarily the distinction between "signals" and "cues." Cues are behaviours or properties of other organisms (more broadly—environmental information) which may influence the behaviour of the potential "receiver", but are not instances of communication. For instance, scent marks left by animals that predators use to track them are cues of prey presence, rather than a signal marking the prey location. In those instances there is influence or information involved, but it is not "sent", nor adaptive to the "sender".

In his treatment, Scott-Phillips argues that while communication may involve information transfer, information is not definitional. He claims that most informational accounts rely on an "everyday" sense of information, which is not precise enough. For him, this reference to information is too liberal to distinguish between "signals" and "cues". However, the application of the most established formal account of information—Claude Shannon's mathematical theory of communication (Shannon and Weaver 1964)—does not help in that regard, in his view.

Shannon's formal framework defines information in terms of reduction of uncertainty. As Shannon was interested in a host of engineering problems related to (primarily) telephony, he assumed a particular model of communication. According to that model, information transfer involves a *sender* which uses a *transmitter* to encode a *message* and send a *signal* over a *communication* channel, which then can be picked out by a *receiver*, decoding the message before it arrives at its ultimate *destination*. He proposed to define the process of communication as a stochastic decision task, where the receiver needs to pick a message from a set of possible messages, conditioned on the signal they have received. In that framework, the "content" of the messages needs to be abstracted, but in turn, communication can be modelled with probability theory. Resultingly, "Shannon information" as the concept has come to be known, corresponds to how much uncertainty the receiver has about the message, prior to the reception of the signal—and how much that uncertainty is reduced (I explore this framework in greater detail in chapter 4).

For Scott-Phillips, Shannon information's exclusion of the contents—what is often referred to as the "syntactic" character of the concept—precludes its usability for defining communication. Indeed, he claims (Scott-Phillips 2008, p. 392) that informational definitions "do not specify how or why a signal is relevant, and hence they have no consistent means by which to distinguish between signals, cues and coercion." And accounting for what are the relevant uncertainties—the contents of the message—requires an "inferential" (in the sense of Stegmann 2016) or adaptationist perspective.[2] Hence, Scott-Phillips (2008, p. 388) eventually opts for a view of communication which sees it as a process of "the completion of corresponding signals and responses", where signal is:

> **Signal**: "Any act or structure that (i) affects the behaviour of other organisms; (ii) evolved because of those effects; and (iii) which is effective because the effect (the response) has evolved to be affected by the act or structure",

and response is, symmetrically:

> **Response**: "Any act or structure that (i) is the effect of some act of structure of another organism; (ii) evolved to be affected by that act or structure; and (iii) which is affected because the other act or structure (the signal) has evolved to affect this act or structure."

Scott-Phillips considers how the adaptationist view may answer the standard counterexamples at length, highlighting that it faces fewer serious problems than the informational alternative: he discusses how adaptationist definition leaves the possibility of unreceived messages and allows for the distinction of non-communicative reciprocal and coercive behaviours. While his model became a reference point in animal communication research, there are some further refinements in philosophical debates. Most interestingly, in the paper by Ramiro Frick and colleagues (2019; see also Bich and Frick 2018), the authors attempt to replace the neo-Darwinian perspective of Scott-Phillips' original formulation (a theoretical approach which has garnered significant criticisms, e.g., Gould and Lewontin 1979) with their preferred organizational framework (e.g., A. Moreno and Mossio 2015).

The organizational approach to biology (A. Moreno and Mossio 2015) focuses on characterising biological systems through their internal organization, which allows them to actively sustain their existence—the property known as *organizational closure* (this point is closely related to, though distinct from, the autopoietic approach, see Maturana and Varela 1980; and the interactivist model, see Bickhard 2009; the organization approach also motivates one of the views on semantic information discussed in chapter 5, see Kolchinsky and Wolpert 2018). While the dominant view of biological function focuses on their evolutionary history, the organizational account describes functions as causal relations within the system which contribute to the maintenance of the system's organization (see

---

[2] This point is also strongly formulated by Owren, Rendall, and Ryan (2010); and Rendall, Owren, and Ryan (2009). These authors argue primarily that the informational approach is underspecified and incompatible with evolutionary principles.

Mossio, Saborido, and A. Moreno 2009). In the context of communication, this notion of function can replace an adaptationist one, which has been central for Scott-Phillips. Effectively, they highlight the "influence-based" nature of signals, and offer to view them more broadly—not only through their evolutionary history, as is the case with Scott-Phillips' adaptationist definition, but also through the immediate contributions of the signals to the organization of the biological systems involved—including in the case of cell-level communication. Nevertheless, their account maintains the distinction introduced by Scott-Phillips (2008). The two approaches to defining communication can be formulated as follows (the wording here follows Frick, Bich, and A. Moreno 2019, p. 2):

> **Informational**: COMMUNICATION is the process of information transfer from a sender to a receiver by means of a signal.
>
> **Influence**: COMMUNICATION consists in functional influence which a sender exerts on a receiver by means of a signal.

Importantly, the wide adoption of the distinction into influence-based and informational definitions ignores the fact that informational perspective is pervasive in scientific literature (see the empirical results discussed in chapter 3 and 4, as well as the theoretical discussion in subsection 1.3.3 below). Further, Scott-Phillips' analysis of the notion of "information" is overly simplistic and ignores crucial complexities involved in the term (I cover this topic at length in chapter 4). Most importantly, information always has physical basis, and the interpretation of the "informational" perspective on communication cannot overlook this fact, even if the use of informational terms in the scientific literature often does abstract from those physical aspects.

Effectively, information is always influence (e.g., consider e.g., the "instructional" notion of information developed by Floridi (2011); this is discussed at length in chapter 5), and this collapses Scott-Phillips' distinction. Especially as informational definitions which appropriately distinguish signals from cues do not rely on information alone.

A more careful reading of information-theoretic literature doesn't allow for such a strict delineation of the influence- and information-based accounts of communication, as Scott-Phillips postulates. Especially as informational definitions which appropriately distinguish signals from cues do not rely on information alone. Indeed, in this thesis I lay out a view of biological communication which notes that neither the functional involvement of signals, nor their informational character can be treated in separation. This view agrees with some existing accounts in the scientific literature (though perhaps not as influential as the Maynard Smith's views on the matter), as will become clear in chapter 3.

Such a conclusion is partially enabled by the fact that both definitions considered by Scott-Phillips broadly follow the same general schema: they both identify a separate sender and receiver, they both have a notion of a vehicle—"signal" or "behaviour"—and they both presume that the communicative system is somehow structured, so that the transfer of the vehicle is possible. Indeed, based on both of those definitions I propose to derive a general definition-schema that will guide my exploration of the literature in the following chapters:

> **Definition-schema**: COMMUNICATION is a process that involves a transfer or exchange of a certain mark K between distinguishable entities, X and Y, enabled by a structure S.

This account is broad, and without the additional specifications—whether we regard the mark K and the structure S in informational, as signal and communication channel, or adaptational terms, as influence and its evolved adaptive traits—does not aid much in understanding communication. Nevertheless, it highlights a particular causal pattern of communication that is confirmed by the corpus analysis presented in chapter 3, and that can be further specified, regardless of the particular theoretical framing of the components included in the schema.

1.3 BIOLOGICAL INFORMATION: GENES, BRAINS, AND SIGNALLING GAMES

While the role that information plays in communication may not be straightforward and remains a subject of debate, the literature on biological communication often intersects with philosophical treatments of biological information. Indeed, there's a number of closely-related problems that crop up in both contexts—chief among them being the naturalistic account of meaning. Throughout the dissertation, I will be connecting these different threads, drawing on various philosophical debates in order to provide a comprehensive account of biological communication. Hence, it is useful to revisit some literature on biological information that informs much of the considerations in this thesis, but works mostly "in the background" of the current debate. These discussions impact significantly the way I frame the issue in the current analysis, hence in this section I will recall them in some detail—reviewing the three main areas where the concept of information in philosophical accounts of biology has become the most relevant.

1.3.1 *Genetic coding*

One of the most widely recognized debates about the role of "information" in biology was concerned with understanding what geneticists mean, when they talk about "genetic information". The debate was a response to the work of Susan Oyama [(2000); the first edition was published in 1985] and was sparked around 1996 with a paper by Sahotra Sarkar. This debate was in fact primarily motivated by the question of what is the causal role of genes in development. In this context, the received wisdom at the time was the claim of "molecular Weismannism" (see Griesemer 2005), also known as the "central dogma" of molecular biology, usually specified in informational terms: the genetic information flows (only and always) unidirectionally, from the genotype to the phenotype. This standard claim is motivated by Weismann's assumption that germ-line and somatic cells are segregated, with only the material in the former being passed on to the future generations. Effectively, it is the statement of the mechanism that ensures that acquired characteristics cannot be inherited. By 1996, the informational formulation, however, pushed the original "material" version to the background: the genetic causal pathway was taken as critically different from e.g., the environmental causal contributions to development, *because* genes contain information about the phenotype, and it is in virtue of these informational properties that they cause phenotypes. The alternative view, emerging from the work within evo-devo (e.g., Oyama 2000), argued for the "causal parity" of the different pathways contributing to development (e.g., Griffiths and Gray 1994).

However, the notion of "information" at play was at the time far from clear. This, paired with the growing skepticism with regard to the causal uniqueness of genes and the arguments for causal parity—the rediscovery of epigenetic effects, the more complex views of development emerging from evo-devo and developmental systems theory, among others—led to growing doubt about the usefulness of informational framing for biological research. The ensuing philosophical debate began as an attempt to analyse the concept and highlight its role in the causal distinction of the genetic pathway. Virtually all participants agreed on several characteristics of the (purported) genetic information:

1. The Shannon, covariance notion of information is too weak (and resultingly, too broad) to be of epistemic use in this context. Instead, the information at play is semantic.
2. The DNA is ascribed informational content because of its arbitrariness: the chemical bonds between genes and their products are contingent, a "frozen" accident.
3. Genes carry information about the proteins that are synthesized on their basis, not about the phenotypes.

But even if all participants agreed on these points, they drew wildly different conclusions from them. Sarkar (1996; 2000), who first noted this reliance of information, staunchly rejected that this notion of "information" can bring any epistemic benefits in biology: initially, he rejected it as a metaphor and not a clear, technical notion, while later he rescinded his criticism slightly focusing on the fact that the assumed characteristics of the informational causal pathway are not in fact unique to genes, and as such this notion doesn't succeed in its expected theoretical role.

The key reference point for the proponents of the informational language has quickly become the paper by the mathematical biologist, John Maynard Smith (2000). Maynard Smith offered the defence of the notion of information, by attempting a technical characterization of the concept that underscored all the key points listed above and added a fourth condition to naturalize the notion of "meaning" at play—in a broadly teleosemantic view, for Maynard Smith the proteins produced from the DNA have meanings because those proteins have a particular function in the cell, for which they were selected in the course of evolution.

Notably, this argument is embedded in a strongly adaptationist view, despite the earlier criticisms of such approach. For instance, in a famous paper, Stephen J. Gould and Richard C. Lewontin (1979) highlighted the inability of adaptationist perspective to capture a number of evolutionary phenomena, for example the various by-products of other biological processes which have not been selected for any particular function, but rather emerged in virtue of the necessary "architectural" constraints of the other process. Effectively, these authors put down a strong argument for a pluralist view on evolutionary change, which suggests that the relevant notion of FUNCTION may in fact require the acceptance of a broader range of causal contributions (more on that in chapter 5).

These two positions have been referenced in a series of papers: Peter Godfrey-Smith (2000; see also Godfrey-Smith 2007) argued in line with Maynard Smith, although he proposed that due to the semantic requirements, the notion of information would be better replaced by that of "coding". His main point is based on a thought experiment, motivated by one of the important hypothesis about the molecular structure of genes (widely discussed before the discovery of the double helix): a protein-template view, where chromosomes would consist of samples of all required proteins that would then be duplicated as necessary. According to Godfrey-Smith's analysis, in this alternative view there would be no coding, no semantic information, as regarding protein-genes in terms of samples would be more direct.

On the other hand, Paul Griffiths (2001) early on has staunchly rejected the use of informational notions in genetics. He argued that the only correct element of that account is the claim of correspondence between the coding fragments of DNA and the sequences of amino acids in the primary structure of proteins. But his main point was more concerned with the uniqueness of the genetic pathway: he argued for the "causal parity thesis," which stated that "[a]ny defensible definition of information in developmental biology is equally applicable to genetic and non-genetic factors in development" (Griffiths 2001, p. 396). In his view, then, it is ultimately of little importance if a notion of biological information applies to genes as the job it is supposed to do (i.e., the distinguishing of the causal uniqueness of genes in development) is pointless (as the genes are not, in fact, causally unique in any important sense). Nevertheless, in more recent work, Griffiths (2016; Stotz and Griffiths 2017) has come around and accepts some use of informational notions in reference to genes. He distinguishes "proximate" and "ultimate" information (in the sense of Mayr 1961). "Proximate" biological information is concerned with how developmental systems use information provided by the genes. In this context, he offered an account of causal specificity (see Woodward 2010) based on mutual information. This view fits the Philip Kitcher's "causal democracy" principle (Kitcher 2003), a version of the "causal parity" view: without privileging the genetic causal pathway, an interventionist account of causation (explained in greater detail in subsection 3.5.1) is able to highlight that genes do carry information for development. For "ultimate" information Griffiths builds on Nick Shea's work (Shea (2007b); Shea (2011); as discussed in chapter 5). However, he replaces the teleosemantic understanding of function proposed by Shea, and instead offers an ahistorical alternative, embedded in Nikolaas Tinbergen's "four questions" framework (Tinbergen 1963), where function is understood in terms of contribution to the "survival value" of the organism.

Kim Sterelny (2000) has presented a similar, albeit more neutral view than Griffiths in his earlier work. He argues that Maynard Smith's account of information offers a coherent view of how it is used if an additional condition—the presence of an "information reader" in the processes that involve information, e.g., in the form of the developmental system—is added. Then, according to Sterelny, the "biological information" view offers a verifiable research program that is ultimately at odds with the Developmental Systems Theory (e.g., Oyama 2000), preferred by Sterelny himself. The key difference in the Developmental

Systems Theory treatment of such cases would be exactly the rejection of the information and reader distinction: any information that can be found, appears in the developmental process and exists only within that process, making the informational lens superficial. This leads Sterelny to a version of "parity thesis" though he underscores that "neither all nor only genes are information carriers" (Sterelny 2000, p. 201), if semantic information even exists in biological processes.

Eva Jablonka (Jablonka 2002) argues for the parity thesis, but differently than Griffiths and Sterelny, she views informational notions as epistemically beneficial. She suggests an alternative definition of biological information: "a source [i.e., an input or a cue, rather than Shannonian sender] becomes an informational input when an interpreting receiver can react to the form of the source (and variations in this form) in a functional manner" (Jablonka 2002, p. 578). The form in this definition relates to the organization or structure of the "source" (as examples of sources, she lists DNA or train timetables), and the relevant notion of function is slightly different from the more popular, teleosemantic one (which will be discussed in detail in the chapter 5): it is defined in terms of consistent causal roles within natural-selection-designed systems that contribute to the goal-oriented behaviours of those systems (Jablonka 2002, p. 582). This notion of information is meant to include the genetic and other inheritance channels (she lists: epigenetic, behavioural, or cultural-symbolic) as well as other instances of biological information transmission, including animal communication.

Slightly later, ending this "first wave" of the debate, in 2005 several authors offered further criticisms of biological information: James Griesemer (2005) views this notion as a "purely functionalist 'generalization-by-abstraction' " which misses key facts about the causal structure of the development and in his view cannot account for the multiple levels of selection at which evolution operates (see Okasha 2008). Marcel Weber (2005) criticizes the use of semantic information in this context as an instance of intentional terms (he lists also "programming", "orchestrating", or "directing" as they appear in the description of the role of genes). Differently from many other critics of the informational view, he rejects the parity thesis as well, and defends a view of genes as "causally privileged" in at least a pragmatic sense. Ulrich Stegmann (2005; see also Stegmann 2014) offers a treatment of informational content of genes which limits the scope of the informational view. He argues that genes contain instructional content about their immediate products, and act as templates (reminiscent of Godfrey-Smith's (2000) thought experiment—though in that earlier analysis Stegmann's proposal would suggest a non-informational character of genes).

What these early discussions shared was the rejection of Shannon information as not epistemically useful for the evolutionary theory—even though all these authors agree that genes carry Shannon information. However, this focus on Shannon information was motivated by a relatively narrow view of information theory, and involved several misconceptions regarding the theory—a topic I explore at length in chapter 4. The recognition of these misconceptions motivated philosophers in what can be identified as a “second wave” of the debate, when few years later the topic of genetic information has returned.

In a paper initially circulated in 2009, Carl Bergstrom and Martin Rosvall (published as Bergstrom and Rosvall 2011), point out the misconceptions regarding Shannon theory of information which plagued the early debate recounted above. They centre the development of efficient coding algorithms that was of critical importance to the engineering side of Shannon’s work (at the same time, Bergstrom and Rosvall frame this problem in terms of “compression”, implicitly referencing Algorithmic Information Theory though they do not explore this further—see chapter 4). In their view, the previous discussions of genetic information have overlooked the transmission model Shannon’s theory presupposes. The processes of transcription and translation that were referenced by the previous participants of the debate are in fact components of the decoder. Instead, the relevant notion of Shannon information is concerned with the transmission across generations, where a sender (the parent) and a receiver (the child) can be actually identified (see chapter 4 for a more detailed discussion of the sender-receiver model and its role in biology). Bergstrom and Rosvall’s detailed analysis of information theory highlights further misconceptions (e.g., the symmetry of mutual information that Godfrey-Smith (2007) pointed out as an argument against the applicability of Shanon theory). Eventually they arrive at a “transmission view of information” which adapts Shannon to the genetics case: “An object X conveys information if the function of X is to reduce, by virtue of its sequence properties, uncertainty on the part of an agent who observes X” (Bergstrom and Rosvall 2011, p. 165). Their central point is that the transmission view is able to do the explanatory work in genetics and development that previous authors (mainly Maynard Smith and Godfrey-Smith) expected from the semantic view, with fewer assumptions required (which some theoreticians, for instance Griesemer, as discussed above, saw as highly problematic).

In the commentaries (Godfrey-Smith 2011; Shea 2011; Maclaurin 2011) that accompanied the paper, several points from the previous stage of the debate have been reiterated. Shea (2011), however, engages the most with the transmission proposal and in his analysis highlights that the view Bergstrom and Rosvall advance is in fact complementary with the older, semantic account. He goes back to Sterelny’s (2000) point regarding information readers, and highlights his own earlier analysis (Shea 2007b) that identifies the components of the transmission model in how genetic information is used by the developmental system.

Slightly later, in a similar vein, Oliver Lean (2014) also proposed that “pure” Shannon information is capable of doing much of the explanatory work expected from the semantic account of the term. He makes two points: first, he distinguishes the information content (in terms of mutual information) from information value, which is viewed in a decision-theoretic perspective (in line with earlier proposals co-authored by Bergstrom, see e.g., Bergstrom and Lachmann 2004; Donaldson-Matasci, Bergstrom, and Lachmann 2010). Second, he highlights that an etiological account of function (discussed in detail in chapter 5) can be used to ground a notion of informational function of some biological systems, without immediately postulating a meaning of that information. A similar point is put forward by Stephen Mann (2020), who highlights that if a Shannonian transmission model is properly applied, the information in the signal becomes relative to the receiver: he cashes out this relationship precisely in terms of a function.

Quite a different route is explored by Brett Calcott (2014) and Ronald Planer (2014) who draw on Brain Skyrms' work on signalling systems' evolution (described in detail in subsection 4.1.4) to account for gene regulatory networks (in the case of Calcott) and the notion of a "genetic program" more broadly (for Planer). Lean, in discussing this line of research, concludes that his model is compatible with the sender-receiver approaches. Similar arguments are also put forward by Godfrey-Smith (2013; 2014), who applies the signalling games model more broadly to biological information, but includes explicitly genetic information as a "memory system with no sender" (I come back to his views in the discussion of animal communication below, subsection 1.3.3). This work aligns with the direction of Bergstrom and Rosvall's earlier proposal, in that the authors, in positing signalling games scenarios for these biological processes, are forced to explicitly analyse instances of genetic coding through the Shannonian communication model (even if their work is conceptual and does not simulate the postulated signalling games).

An alternative comes from the established, though relatively absent from biological debates, account of information as well-formed, meaningful data, proposed by Luciano Floridi (2010; 2011). Floridi focuses on the notion of "information" in the everyday, epistemic meaning—he claims that "the term 'information theory' [for Shannon's mathematical theory of communication] is an"appealing but unfortunate label, which continues to cause endless misunderstandings" (Floridi 2010, ch. 3). In his terminology Shannon "information" is considered data: the patterns or fields of differences, instances of a lack of uniformity in the world. If such data are well-formed (structured) and meaningful, they are considered information, as Floridi's formula expresses: "information = data + meaning." His usage sometimes waivers—for instance, he refers to "environmental information" (what is called "natural signs" in more common terms, e.g., rings in the wood of a tree as a sign of its age), though such information is not always meaningful, and hence should rather be considered "data." Yet, his approach offers an intriguing solution to the problems related to genetic information, overlooked in the above discussed debate. Floridi himself considers biological information only briefly (Floridi 2010, ch. 6) and proposes to take it as instructional environmental information. In his views, DNA *is* information (rather than "contains" or "carries" it). Such information is instructions which "contribute to control and guide the development of organisms" (Floridi 2010, p. 80). As such it can be understood (partially) through the mathematical theory of communication, but it is "information *for* something, not *about* something" (Floridi 2010, p. 80)—it lacks reference, and as such doesn't require (or allow) the use of a semantic lens. This view has been recently explored further by Thomas Dickins (2023). Dickins finesses Floridi's account, offering to view the requirement of "meaningfulness" as a "semantic context". This notion highlights that data becomes information for some system, when it is accepted (as input) by that system within a particular semantic context. Context itself emerges from the system's design. In the case of genetic information, the semantic context is precisely that of DNA acting as instructional with regard to the development. Dickins then proceeds to uncover traces of such a notion of instructional information in the Maynard Smith's account (2000, and more broadly within the modern synthesis view), arguing that if we take this perspective, and at the same time regard the notion of genetic information as an idealization (as the proponents of Modern Synthesis did, in his view), the challenge from developmental theorists and their insistence on causal parity can be easily dealt with.

To my knowledge the only other account explicitly going beyond the standard, Shannonian notion of information—in this case, adopting the notion of information content offered by Algorithmic Information Theory (see the discussion in chapter 4)—is proposed by Arnaud Pocheville (2018). He casts what I term the "transmission/content distinction" (in chapter 4) as a "choice and construction" distinction. "Information as choice" (i.e., transmission) is concerned with "choos[ing] between a set of well-defined alternatives with a well-defined probability distribution" (Pocheville 2018, Section 2)—e.g., in genes guiding the translation and transciption into proteins. In this context, Pocheville accepts the more recent Griffiths' account of causal specificity of genetic information. His own contribution focuses on accounting for "information as construction", i.e., information content of DNA. For this purpose he adopts the notion of Kolmogorov complexity (see subsection 4.1.3) and analyses in detail information available in genes, and then connects it with a proposed description of operations of the biological cells, formulated in terms of a (programming) language of a cell. Setting aside the debate about the accuracy of the analogy, it provides him a framework that can capture the information stored in the DNA sequence *for the cell*, as the operations of the purported programming language, such as "transcribe(sequence)" or "translate(sequence)". In this work Pocheville takes advantage of the overlooked formal capacities of Algorithmic Information Theory and is able to posit a clearly realist model of genetic information, one that escapes much of the existing criticism that has accrued in relation to older approaches (regardless of whether those criticisms are well-founded or not, as for instance Weber's (2005) total rejection of intentional framing, which likely significantly misses the mark).

1.3.2 *Neural coding*

The central role of "information talk" (e.g., Griffiths 2001, p. 394) in genetics can only be compared to the significance the concept has in the contemporary study of cognition—across psychology, cognitive science, and neuroscience. It has become commonplace to talk about cognition as "information processing," and in fact the origins of this perspective can be pinpointed relatively precisely in 1956 (see G. A. Miller 2003; Newell and Simon 1972, p. 4), as the beginning of what has come to be known as the "cognitive revolution". Besides introspective statements of the figures involved in the early history of cognitive science(s), historical research confirms these points. Margaret Boden (2006, p. xxxv) in her expansive account of the emergence and development of cognitive science lists three central characteristics that in her view are decisive for the unity of this field: the focus on the mind (mental processes, mind's relation to brain) as the explanatory target, the role of "information answers" to the questions of the field, "whether to recommend them or to criticize them", and the conceptual framework of the mind as a machine—with all the theoretical tensions it brings.

As Boden shows, the application of Shannon's information theory to psychology was largely motivated by its formal character—as the field at the time sought for mathematical ways of describing cognitive operations to overcome the limitations of strict behaviourism. Information was quickly adopted as a measurement tool, e.g., to describe the channel capacity of the different sensory modalities, and then quickly expanded in more theoretical directions. She credits George Miller as an early adopter of information theory in psychological context, responsible for its eventual theoretical role in the nascent cognitive science research program. In that narrative (a view that's quite widely shared), Miller's famous paper "The Magical Number Seven, Plus or Minus Two" (1956) is the point of introduction of this theoretical function of information language: e.g., he explains the apparent limits on the memory capacity by pointing to the notion of channel capacity, and considers "recoding by chunking" as a mechanism that allows to overcome these limitations. The application of information theory to memory was quickly followed by the use of the framework to think about consciousness (in the sense of access consciousness, see Block 1995), e.g., to account for attention and its selective character (e.g., Broadbent 1952). These developments were soon followed by the application of the conceptual framework of computation (Boden credits Jerome Bruner, and especially Bruner, Goodnow, and Austin 1956), which eventually covered all central areas of research in psychology, including language (e.g., Chomsky 1959), problem solving (e.g., Newell, Shaw, and Simon 1958; Newell and Simon 1972), or perception (Marr [2010]1982) and eventually shaped much of how we think about cognition.

Despite the wide adoption, the understanding of cognition in terms of information processing—and especially in terms of computation—remains a topic of debate for philosophers. While it is likely impossible to eliminate the language of information from discussions of the mental, even if one would like to, there are genuine theoretical and philosophical concerns over the implications of this perspective. For instance, the early version of computationalism in cognitive science promoted a view of cognition as purely functionalist, disembodied process which soon led to a number of criticisms (e.g., Dreyfus 1986; Dreyfus 1992; Van Gelder 1995; see the overview in Rowlands 2010) and a host of alternative proposals, broadly termed the "4E" framework for the various E-labels these approaches use:[3] ecological (Gibson 1966, 1986), embodied (e.g., Barsalou 1999; Brooks 1991; Gallagher 2005; Lakoff and M. Johnson 1980; Merleau-Ponty 2005), enacted (e.g., Di Paolo, Buhrmann, and Barandiaran 2017; Di Paolo, Cuffari, and De Jaegher 2018; Maturana and Varela 1980; Thompson 2010; Varela 1986; Varela, Thompson, and Rosch 1991), embedded (e.g., Hutchins 2006; Sterelny 2010), extended (e.g., Clark and Chalmers 1998; Clark 2008), as well as (less frequently) emotional (e.g., Damasio 2000; M. Johnson 2010) or evolved (e.g., Barrett 2018). These various perspectives commonly rely on the formalism of dynamical systems theory (e.g., Thelen and L. B. Smith 1994; Van Gelder 1998; Chemero 2009) and sometimes draw on the phenomenological account of the mind (most notably, Merleau-Ponty 2005). The rejection of computationalism in some of those approaches is associated with a strong opposition to "representationalist" explanations of cognitive processes (e.g., Chemero 2009; Van Gelder 1995)—i.e., explanations which require the existence of mentally instantiated stand-ins for external objects, sometimes symbolic (a strong version of representationalism is the language of thought hypothesis, Fodor 1975), which are variously processed (hence the relation to computationalism). Nevertheless, some anti-representational 4E approaches do accept the informational perspective of cognition—to mention only the concept of "ecological information" used by James J. Gibson in his treatment of optical flow (1966; 1986).

Indeed, there is a broad sense of "information" (mostly in the syntactic, correlational sense captured by Shannon's formalism) that is uncontentious in contemporary study of cognition. Methods of information theory are used widely across the cognitive sciences, to describe cognition both at the subpersonal level of neuronal operations, and the personal level of cognitive processing (e.g., Arbib 2003; Sayood 2018). Nevertheless, the use of information language of "coding", with its representational implications and role in computational accounts of mind (e.g., Haugeland 1993; Miłkowski 2013) has been subject to scrutiny. While many debates are concerned with the semantics of information and will be discussed in detail in chapter 5, some preliminary discussion is in order.

The origins of "coding" in neuroscience can be traced back to 1960s (e.g., Barlow 2012; Perkel and Bullock 1968; for contemporary review, see the article "Neural Coding" and related entries in Arbib 2003). Coding in this context is understood as "the transformations of information in the nervous system, from receptors through internuncials to motor neurons to effectors" (Perkel and Bullock 1968, p. 307), but it can be analysed into several related concepts which differ in their assumptions. Romain Brette (Brette 2019, pp. 1-2; see also the criticisms in Bickhard and Terveen 1996) identifies three main senses:

---

[3] The number and exact choice of E-cognitions to include is a matter of debate among the proponents of this approach, as the various labels overlap in different ways, and depending on one's priorities and theoretical preferences some appear better-suited than others. Hence, there is more than 4, and not only E-, labels I list here as the approaches constituting this general framework. The standard selection is embodied, embedded, enacted, and extended.

- "correspondence" between two domains (e.g., intensity of a visual signal and the rate of a resulting spike train)—this is the level which is best understood in terms of Shannon information,
- "representation", since the spike trains are stand-ins or messages for the receiver—introducing semantics into the message,
- and "causality", which must be present for the representation to occur–the signal must have a causal effect on the receiver, e.g. the visual stimuli are encoded by the spike trains they cause, not the other way around, even though Shannonian information terms are often symmetrical.

Brette uses this analysis to argue against the adequacy of "coding" in the neuroscientific context. He notes that the majority of the literature assumes the "representational" or "causal" sense of coding, but employs the tools of the "correspondence" view—which are not powerful enough to capture the representational sense. For instance, in terms of the temporalities of processes they allow, the correspondence perspective is indeed highly misleading: "substituting the arbitrary temporality of algorithms for the [continuous] temporality of the underlying physical system", Brette (2019), p. 2.

While scientists' use of the term at times overlooks the limitations Brette points out, philosophers have realized them early on. One of the earliest and most influential attempts to go beyond the "correspondence" sense in relation to mental representations is the causal-informational account offered by Fred Dretske (1981; 1995). Dretske's account is centred on the notion of "information-carrying" or "indicating": a content of a mental representation is what that representation has the function to indicate. The "indication" relation is defined as follows: "an event token of type R indicates that a state of affairs of type C~obtains if and only if (i) C in fact obtains and (ii) the probability of C's obtaining, given that R is instanced, is one" (Schulte and Neander 2022; cf. Dretske 1981, p. 65). This co-occurrence cannot be merely incidental, and often is underlay by a causal connection between C and R—though this is not a requirement. While there's a number of problems with Dretske's account (covered in greater detail in chapter 5), he established a general framework that continues to inform the debates in this area.

Indeed, philosophical analysis of the concept of "representation" and its role for explaining mental function has attempted to bridge the gap between the "correspondence" and "representational" coding (in Brette's sense) in various ways (for a recent and relatively broad contribution, see the essays in Smortchkova, Dołęga, and Schlicht 2020). This topic has been highly contentious and the "representation wars" (Clark 2016a) have shaped much of the debates in philosophy of mind from the 1990s through 2010s. Recapitulating these discussions in detail would distract from the main issue at hand, as INFORMATION—while present in the background—has not played a key role in those debates. Instead, I will conclude this section looking at a more recent example of the "coding" language, namely the methods of "neural decoding" (based on the analysis in J. B. Ritchie, Kaplan, and Klein 2019; see the review in Haxby, Connolly, and Guntupalli 2014), as they allow for a relatively brief introduction of several problems recurring throughout the thesis.

"Neural decoding" is the label used for a number of methods in neuroscience, in particular systems neuroscience, for "break[ing] the neural code" (Haxby, Connolly, and Guntupalli 2014, p. 435). More strictly, these methods aim at "extracting information from measured activity [to] afford understanding of how percepts, memories, thought, and knowledge are represented in patterns of brain activity" (Haxby, Connolly, and Guntupalli 2014, p. 435) across different species, including humans. Decoding starts with measurements conducted using methods such as single cell recordings or electrocorticography (ECoG), electro- and magnetoencephalography (EEG and MEG) and functional magnetic resonance imaging (fMRI—most popular in this context, due to its non-invasive nature and high spatial resolution). These data is then processed through a number of related algorithms that aim to differentiate patterns of activity in relation to the stimuli. There are three main approaches in this context (see Haxby, Connolly, and Guntupalli 2014): multivariate pattern analysis (MVPA, originated by Haxby et al. 2001), representational similarity analysis (RSA, see Kriegeskorte 2008; Kriegeskorte and Kievit 2013) and stimulus-model-based encoding and decoding (Naselaris et al. 2011) which recently grew in popularity partly in virtue of the application of deep learning methods (e.g., convolutional neural networks, see Shen et al. 2019; machine learning methods can in fact be applied across the different decoding "tasks", see Glaser et al. 2020). These approaches differ in the exact algorithms used. For example, MVPA relies heavily on linear classifiers which aim to cluster different patterns of activity in a multidimensional space (corresponding to the number of dimensions of data—e.g., number of different voxels in a region of interest recorded with fMRI), RSA uses representation dissimilarity matrices (matrices of correlation distance between different stimuli or conditions), while stimulus-model-based decoding originally relied on Bayesian techniques, such as the maximum a posteriori estimation or Kalman filters (see Kalman 1960), and more recently applies different neural network architectures (see also Lindsay 2021). Further, they differ in the insights into brain activity they allow (following Haxby, Connolly, and Guntupalli 2014, p. 437): MVPA methods can reliably distinguish brain states and characterize their organization, RSA uncovers the representational structure (geometry) of brain activity (e.g., can show that the representation of the same stimulus in different regions have different structures, even if MVPA classification accuracy is equivalent in those regions), while stimulus-model-based encoding and decoding can relate brain activity patterns to "the constituent features of stimuli or cognitive states" (Haxby, Connolly, and Guntupalli 2014, p. 437).

J. Brendan Ritchie and colleagues (2019) note that the decoding methods (they group them together under the label MVPA, but discuss the different applications listed above) rely on an inference pattern they call the decoder's dictum:

> If information can be decoded from patterns of neural activity, then this provides strong evidence about what information those patterns represent. (J. B. Ritchie, Kaplan, and Klein 2019, p. 582)

The decoder's dictum is concerned precisely with the relation between information and representation, or correspondence and representational sense of coding, to use Brette's distinction. Crucially, it rests on the assumption that neuronal population codes involve linear transformations, and hence, MVPA methods which also rely on linear algorithms are biologically plausible and sufficiently similar in architecture to brain's own decoding processes, so that if significant decoding is achieved it can be inferred that the information is in fact directly available for neural processing (J. B. Ritchie, Kaplan, and Klein 2019, pp. 592-593).

However, as Ritchie and collaborators claim, the dictum is false: in many instances it is still unclear why the applications of decoding work—for example, MVPA on fMRI data from primary visual cortex (V1) is able to decode information regarding stimulus orientation despite significantly lower spatial resolution compared to the size of the cortical columns which are known to encode the orientation in this region. This is a paradigmatic case, highlighting that the information decoding algorithms latch onto need not correspond to information available directly for neural processing, despite the assumptions of the dictum. While these charges are not fatal to the decoding programme, but rather urge a rethinking of how to interpret the results reliably (J. B. Ritchie, Kaplan, and Klein 2019, pp. 599-601), this discussion highlights the central role of "information" and "coding" for neuroscience. Indeed, as Ritchie and colleagues note in passing, this research programme is motivated by the view of the brain as a mechanism encoding and decoding information (e.g., sensory stimuli). The development of decoding tools assumes that this information is present in the measurable activity patterns, and can be accessed. The informational sensitivity of decoding methods is cited as one of its important epistemic benefits (J. B. Ritchie, Kaplan, and Klein 2019, p. 587-588). And the decoder's dictum (even if mistaken) plays a crucial role in research because information directly used in neural processing is seen as a central explanatory target of neuroscience (especially in systems neuroscience, where these methods are particuarly popular—see Kriegeskorte 2008; Kriegeskorte and Kievit 2013).

1.3.3 *Animal communication*

Animal communication has developed into a research area of particular interest to philosophers, who draw from animal behaviour studies to substantiate views in philosophy of language. As discussed above (in section 1.2), while informational framing of communication is contentious, it is persistent in the literature. Colin Allen and Marc Hauser (1992) develop Dretske's notion of information (1981), as well as Millikan's framework (1984) in the context of debates about the internalism and externalism in animal communication, to propose a "strong" notion of informational content. They highlight that informational analysis can help distinguish between signals which refer to an internal state of the sender (e.g., ethylene molecules picked out by acacia trees in result of damage—which do not allow the trees to distinguish between artificial wounding and that inflicted by kudus, which naturally feed on the trees, but rather signal the damage itself) and an external state or object (e.g., rhesus monkey vocalizations—produced only if they see another group member).

There's a number of more recent positions which argue that information is definitional for communication, *contra* the points raised by Scott-Phillips (2008) and Owren, Rendall, and Ryan (2010). For instance, Andrea Scarantino (2013; 2015) and R. Haven Wiley (2013) develop independent arguments that reject the view of incompatibility of information and influence, and argue for an explicitly informational definition of communication that is still sensitive to evolutionary effects. Caitlin Kight and colleagues (2013), in turn, highlight that informational understanding of communication enables a detailed investigation of the processes responsible for the receiver's response—an element relatively overlooked in research—which they propose to model through statistical decision theory. The joint picture emerging from those views is that information and influence are not mutually exclusive. As I will argue in various places throughout the thesis, it is necessary to include the physical basis of information in any account of communication. In result, information *can directly be the causal influence*—a point Luciano Floridi also made in the context of DNA (as referenced above in subsection 1.3.1). If all informatyion is physically instantiated, for the signal to be picked up by the receiver, they must be influenced by that physical carrier (except for potential cases of action at a distance, which are unlikely to be relevant for the scales at which biological systems function).

Most importantly, the informational understanding of communication has motivated a broad research programme which studies communication through signalling games (Skyrms 2010).[4] The general idea of signalling games is to employ game-theoretic concepts (and, more recently, the computational agent-based modelling framework) to model communicative interactions. The simplest versions of these games (Lewis 1969) assume a sender and receiver with asymmetric access to the world (sender sees the stimulus the receiver must react to) and with shared interest (an equal reward is paid to both, conditioned on the receiver's "action"; I cover the signalling games framework in greater detail in subsection 4.1.4). The close relation between signalling and broader evolutionary games allows for explicitly studying informational properties of interactions

---

[4] As a recent review (Zubek, Korbak, and Rączaszek-Leonardi 2024) highlights, the signalling games framework originally, as proposed by David Lewis (Lewis 1969), has not employed the concept of information, and indeed a significant body of research it inspired—particularly in experimental semiotics (e.g., Galantucci 2005; Roberts 2017; Roberts and Galantucci 2016)—does not rely strongly on informational terms. Nevertheless, there are numerous especially computational applications of this framework which do centre on information, and the position the signalling games framework has attained is largely due to those uses.

while preserving the evolutionarily relevant characteristics of those interactions (in direct response to some of the charges against the informational views presented in Owren, Rendall, and Ryan 2010; Rendall, Owren, and Ryan 2009).

This perspective is picked up by Peter Godfrey-Smith (2012; 2013; 2014; Shea, Godfrey-Smith, and R. Cao 2018), who develops the Brian Skyrms' (2010) version of the signalling games framework to cover various instances of biological information: from genes, via memory, to animal communication. The most developed version of the framework is presented in the joint paper with Nicholas Shea (whose views are discussed in detail in chapter 5) and Rosa Cao. They build on Skyrms' views of "informational content" in signalling systems which he views in terms of changes in subjective probabilities of states of world for the receiver, broadly speaking (see also the discussion of Isaac (2019) in subsection 5.2.2; cf. Scarantino (2015)). However, they argue that there is a special case of content exemplified by signalling systems at an evolutionary equilibrium which the Skyrms' notion does not capture—as it is focused on the sender's perspective and disregards receivers' response.[5] They offer the notion of "functional content" intended to supplement this view. A signalling system has functional content only at equilibrium, if the receiver produces a specific behaviour in response to (at least some) signals, such that in the state of the world which correlates with the sending of that signal, the behaviour "produced by the message contributes to the stabilization of those sender and receiver behaviours" (Shea, Godfrey-Smith, and R. Cao 2018, p. 1016).

Signalling games and information are tied more directly in the proposal put forward by Stephen Mann (2018). Crucially, Mann notes that the majority of the literature on information in animal communication views it as descriptive or referential, overlooking the alternative "instructive" (Floridi 2011) or "directive" (Millikan 1995) notion (discussed also in section 5.1). The difference is in direction of fit between the signals and the world. Skyrms' notion of informational content is the example Mann employs: if we consider information as the change of probability of the states of the world, this is a descriptive notion. However, if we understand the content as the change of probabilities of receiver's behavioural responses, the signal acts as an instruction. Mann analyses the honey bees' (*Apis mellifera*) waggle dance (see also section 5.1) through this lens, highlighting that if information content is viewed as instructional, there's a natural notion of the quantity of information emerging from considering the accuracy of receiver bees—which in turn has direct connection to fitness and evolution.

Regardless of these developments, criticisms of informational framing of animal communication persist. E.g., more recently Kalkman (2019) has argued

---

[5] Shea, Godfrey-Smith, and R. Cao (2018) note: "If a sender and receiver have rules configured so that the sender maps states to signals one-to-one, and the receiver maps signals to acts one-to-one, but in a way that guarantees that the act produced is the *worst* one possible in each state, signals have the same informational content they would have if the sender was performing the same mapping of states to signals, but the receiver was producing the *best* act in each state." (p. 1014, emphasis added) Namely, the informational content is a feature of the signal regardless of whether the receiver actually makes use of that content or not.

that neither informational nor influence-based definition are able to adequately capture biological communication, both being too liberal—though Kalkman does concede that information transmission is an undeniable feature of communication. For this reason in the remainder of this thesis I adopt an agnostic approach, as exemplified by the general definition-schema introduced above. Eventually, I will arrive at a view of communication that sees information transmission as a central feature of the process, and I hope that the analysis conducted in chapter 4 and 5 will present this point precisely.

1.4 CAVEATS AND LIMITATIONS

There are a number of topics that the philosophical literature on biological communication and information treats as central, which I nevertheless do not directly engage with here. This includes, most notably, the problems surrounding the possibility of false information: variously classified as misinformation, disinformation, errors, etc. (e.g., Bielecka and Miłkowski 2020; Dretske 1986; Melander 1993). Further, there's a number of considerations in philosophy of language and its intersection with logic and philosophy of mind which I do not include in this thesis (e.g., the Bar-Hillel and Carnap paradox of semantic information, Carnap and Bar-Hillel (1952); or the Scandal of Deduction). This decision was motivated primarily by the focus of the current account on the connection of organism-level biological communication to communicative processes at smaller spatial scales. Even if I'll argue that it is possible to ascribe meaning to signals at those lower scales of biological organization (chapter 5), I do not think that there's a reasonable application of the notion of "truth" at such a scale (but see the discussion regarding the notion of "correspondence" in chapter 5). Nevertheless, as the account I develop here connects closely to frameworks that offer extensive conceptual toolboxes for considering such complex cases, as is the case with many of the teleosemantic theories that inform the discussion of semantics in chapter 5, I do believe that in principle it would be possible to answer such cases within the current proposal.

I also do not cover explicitly a wide range of non-communicative, co-adaptive interactions, which the influence-based accounts of animal communication often consider at length to highlight the distinctiveness of communicative processes (see e.g., Kalkman 2019). However, I consider various edge cases and contrast classes throughout the thesis and I believe that the analytic approach I employ in those cases extends naturally to such standard cases.[6] Again, this choice of focus is motivated by the extensive literature which does include those cases—but overlooks the lower-level processes of signalling that are of central interest here. Since the account offered here is embedded in that literature firmly, again, I believe that a number (depending on particular theoretical choices on which my framework is agnostic) of responses to such arguments can be offered.

Finally, while I draw in various places of the thesis on views which can be placed in the biosemiotic tradition, I do not engage with this literature more thoroughly (for important treatments in biosemiotics see Barbieri 2008; Hoffmeyer

---

[6] One such non-communicative, co-adaptive interaction is the reciprocal grooming, common among the primates. Scarantino (2013) offers this as a counter-example to influence-based theories, as they fail to properly classify this as a non-communicative behaviour: there are two agents engaged, and there's a pair of behaviours (consisting of mutual grooming) which have co-evolved, with the latter act of grooming being a response to the former. Scarantino opts for using the informational definition to properly classify grooming. The framework proposed here, however, introduces the energetic asymmetry criterion (considered in chapter 3 and 5) which leads to proper classification of reciprocal grooming outside the scope of BIOLOGICAL COMMUNICATION.

2008), and only discuss it in some length in the conclusions. This is largely due to the fact that the biosemiotic view of life as based on semiotic processes—signs and codes—presupposes the speculative view of COMMUNICATION as a central characteristic of life which I wish to explicitly bracket from the considerations in chapters 3-5. As will become clear, this is not a major obstacle, especially since—as I will discuss in some detail in chapter 6—the view of communication proposed here is slightly more deflationary, compared to the stronger symbolic requirements that some biosemioticians endorse, and significantly more materialistic than some biosemiotic positions.[7] Further, the understanding of biological communication some proponents of biosemiotics endorse—as "more than just machine-like exchange of information", a process "based on *semiosis*, or sign processes" (Hoffmeyer 2008, p. xiv)—did not come out of the empirical data during analysis. As such, I decided to focus here on more mainstream interpretative frameworks for the majority of the thesis and contrast them with the biosemiotic perspective only in the conclusions.

---

[7] These idealistic assumptions are traced by L. Han (2021). She notes that they derive from Jakob von Uexküll's theory of *Umwelt*, notably closely related to Nazi ideology (see Schnödl and Sprenger 2022), and traces them to the work of Thomas A. Sebeok, one of the earliest proponents of biosemiotics, who himself regarded von Uexküll as the first biosemiotician. For instance, Sebeok (2001) notes: "Without necessarily committing oneself to this or that brand of idealism [...] it is clear that what semiotics is finally all about is the role of mind in the creation of the world or of physical constructs out of a vast and diverse crush of sense impressions" (p. 37). Other biosemioticians, however, reject the opposition of idealism and realism. E.g., John Deely (Deely 2011) notes that "Semiotics begins by replacing the epistemological paradigm of idealism with its own paradigm, that of the sign understood and taken according to its own proper being – relation in its singularity as indifferent to all subjective divisions of being, including the division between what is mind-independent and what is purely objective. Such a being as constitutes the sign is neither that of an object nor that of an idea, still less that of a thing, but quite simply that of a suprasubjective function whereby what the sign is not is presented by the sign to an interpretant as the sign's significate or 'objective content' " (p. 130).

2

DIGITAL PHILOSOPHY OF SCIENCE: A METHODOLOGICAL PRIMER

The twentieth century has seen a sudden growth of popularity of philosophy of science: after the Second World War, between the year 1940 and 1960 the popularity of the term grew over 3,5 times (according to Google Books Ngram corpus search for the use of the phrase "philosophy of science"). This increasing popularity is likely tied to the growth of science itself (see Ankeny et al. 2011), which might have impacted the direction the philosophy of science has been evolving in. Especially in the twenty-first century, we've seen a growing distrust towards the so-called "armchair" philosophy and increasing reliance on various sources of empirical data: a methodology, broadly speaking, modeled after that of natural sciences. Most notably, this includes the "naturalistic turn" in philosophy of science (see Rouse 2023) or the experimental philosophy program in philosophy of language, mind, epistemology and ethics (see Knobe 2007). Digital philosophy of science positions itself as a further development of those approaches (see De Block and Ramsey 2022), which responds to some of the limitations of the previous, case study-based methodology of naturalistic philosophy of science, especially in the face of the rapid growth of the number of scientific publications and science more broadly.

In this chapter I begin in section 2.1 by drawing the theoretical background for the digital methodology, expanding on the general remarks presented in chapter 1. Then, in section 2.2, I briefly review main methodological strands of computational work in philosophy of science: network analysis (most importantly, citation networks) and distributional semantics, which will be the main focus here. After introducing basic concepts and assumptions and reviewing key studies that have paved the way for the current project (section 2.2), I discuss some of the crucial open methodological questions and limitations of the digital approach (section 2.3). Finally, I present the research design of the digital studies conducted for this dissertation and reported in the remainder of this monograph (section 2.4 and 2.5).

2.1 PAYING ATTENTION TO THE PRACTICE OF SCIENCE

Naturalistic philosophy of science (following the brief historical sketch in Rouse 2023; see also the interviews in Callebaut 1993) began in the late 20th century, in response to the growing discontent with limitations of the previously dominant, logical approach. Roughly speaking, beginning with the Vienna Circle programme of logical empiricism, the guiding idea for philosophers of science was to establish logic as the foundation for scientific method (Giere 1985). One of the significant impacts of Thomas Kuhn's *The Structure of Scientific Revolutions* ([1962] 2012) was to advocate for a more central role of history of science for philosophical reflection. Ronald Giere, one of the early proponents of the naturalistic approach (Giere 1985; see Rouse 2023), sees this as an inception point of the approach that seeks to position philosophical study of science firmly in relation to the empirical research in "history, sociology, psychology, anthropology, and feminist studies of scientific work" (Rouse 2023, p. 720). This approach requires philosophers to take the descriptive function of philosophy of science seriously: while they may ultimately pursue normative goals or seek to analyse and clarify central concepts of scientific research, the naturalist perspective requires an alignment with how science is actually carried out. This differs significantly from the logical approach, which took as its reference point some abstract, idealized view of science.

One corollary of this approach is that philosophers of science needed to shift their methodologies to hold accountable to the standards in the various strands of science studies. The immediate solution was to combine standard methods of philosophical analysis (primarily aimed at published outcomes of scientific research) with the results coming from the different disciplines of science studies, primarily history (e.g., Giere 1973, 2011). However, over time the limitations of these methods began to garner a growing attention (e.g., R. Scholl and Räz 2016; Steel, Gonnerman, and O'Rourke 2017), highlighting the constraints of the dominant approach.

These methodological worries were accompanied by a growing realization of "the historical vicissitudes, institutional embeddedness, material realization, normative complexity, disciplinary diversity, and political engagement of scientific understanding in practice" (Rouse 2023, p. 720; to support this claim, Rouse cites, among others, Ankeny and Leonelli 2016; Barad 2007; N. Cartwright 2019; Longino 1990; Nersessian 2010; Winsberg 2018; see also the socially crucial work of Naomi Oreskes, Oreskes and Conway 2010; Oreskes 2021). In result, the focus of philosophy of science shifted increasingly from the relatively stable outputs of science—in the form of "theories" or "scientific knowledge" and their relation to the world—to the malleable and dynamic practices of how science is *actually* carried out.

This focus on practice and its social embedding has been particularly visible in the research field of Science and Technology Studies (STS). While the roots of STS reach back to 1930s and the work of Ludwik Fleck (Fleck [1935] 2008; see

Latour 2005; Jarnicki and Greif 2024), the field took its form with the publication of the highly impactful study of Bruno Latour and Steve Woolgar, *Laboratory Life* (Latour and Woolgar 1979), where the two researchers adapted the anthropological methodology of participant observation (the method has seen one of the earliest and most influential applications in Malinowski 2002; see also the methodological discussion in Geertz 1973) to the study of research practices in a biological laboratory. Latour and Woolgar spent twenty one months at Roger Guillemin's laboratory at the Salk Institute, observing the practices of employed scientists, analysing all the literature—published and unpublished, so-called "grey literature", i.e., drafts, internal communications, data sheets etc. (see Auger 1998)—produced at the laboratory, and interpreting the results (Latour and Woolgar 1979, pp. 39-40). This study, followed by many others, has constructed and documented the view of science as a human activity, with a variety of social, historical, material, economical, and other entanglements—just as other human activities.

The radical social constructivism of the STS work has led to significant push-back from philosophers of science, especially those who did not strongly subscribe to the naturalistic approach. While some philosophers (e.g., Hacking [1983] 2007) admitted the impact of the social dimension of scientific practice, they still positioned the epistemic outcomes of science independent of these circumstances. Nevertheless, a growing number of philosophers of science sought to incorporate STS insights into their work. Eventually, this has led to the appearance of the field of philosophy of science in practice (e.g., Ankeny et al. 2011, and other articles in that issue; Soler et al. 2014). The goal of this approach is to complement the standard questions of philosophy of science regarding the epistemological standing of scientific theories with investigations of actual research practice, often conducted with STS methodology, though without its constructivist assumptions. As Rachel Ankeny and colleagues put it, "[i]f we are interested in exploring the assumptions and methods underlying the sciences, it is essential not only to explore the theories and results produced by scientists, but the processes by which they came to these conclusions" (Ankeny et al. 2011, p. 304), without losing sight of their relation to the world.

In the context of these developments, the application of methodologies of digital humanities is a specific approach to naturalistic philosophy of science. Building on the methods of natural language processing, digital philosophy of science benefits from the advantages of computational tools which allow the researchers to explore a volume of data that before remained outside of the scope of human cognitive abilities.

The standard motivation (see e.g., De Block and Ramsey 2022) for the inclusion of digital tools can be laid out as follows: the reliance of naturalistic approaches on close, careful case studies of scientific work (both practices and epistemic outcomes) has been necessarily restrictive in scope. Often large swaths of literature have relied on the same examples, as philosophers argued over interpretations proposed in extant work (Burian 2001; Pitt 2001; Currie 2015; Steel, Gonnerman, and O'Rourke 2017, one such example is the Lotka-Volterra model, commonly used in the analysis of scientific representations).[1] While this approach is still capable of providing significant insights into the inner workings of science, it has important limitations: for obvious reasons, the examples were taken from the work the philosophers were already familiar with, and hence there was little information available to judge the generality of the conclusions drawn. This is precisely the point that has driven the development and use of digital tools for the study of scientific papers. The achievements of computational linguistics and natural language processing, as well as their uptake in other disciplines, especially history, known under the umbrella name of "digital humanities" (see Berry 2012b), promised philosophers the ability to expand the scope of their work and analyse bodies of literature which go far beyond what can possibly be read during a human lifetime.

At the same time, it is impossible to do philosophy off of only quantitative data. Hence, these methods are most usefully implemented in a more complex, multi-level or mixed methods approaches, which attempt to use computational investigation to establish the generality of the dataset, but supplement them with more narrow, qualitative analysis of individual case studies. These approaches are covered in greater detail in section 2.4.

2.1.1 *Conceptual practice of science*

While the digital methods have gained traction in naturalistic study of science, they have not received as wide attention in the study of scientific practice. This is primarily due to the focus of digital methods on the textual outputs of scientific work. Such outputs, first, put emphasis on the epistemic outcomes of science, not on the practice, and, second, present a distorted view of what *actually* happens in the process of scientific investigation (e.g., Rouse 1990; Schickore 2008). Indeed, the most direct input of digital methods is for the study of scientific knowledge (or more precisely, how scientific knowledge is presented; e.g., Overton (2013); in one of the first digital studies, focused on the role of explanation for scientists), and when the dimension of practice is considered, the methods pose significantly more difficulties. As an example, consider a recent study by Christophe Malaterre and Francis Lareau (2024a; 2024b), where the authors attempt to infer and reconstruct social structure of particular research

---

[1] A recent study (Mizrahi 2020) aimed to quantify this prevalence of the method of case studies. While it appears to confirm the earlier anecdotal observations, its limited methodology decreases credence in the results. The study relies on a simple keyword search for "case study" and/or "case studies", using the percentage of philosophy of science articles (defined through the location of publication, either the PhilSci-Archive or some manually selected journals, such as the British Journal for the Philosophy of Science) returned by the search as a measure. This classifies any mention of "case study" as the reliance on that methodology, which is too permissive. Hence, the results of the study have limited informativeness. Note that this study follows the tendency of some of the digital studies to simply reaffirm existing knowledge, a criticism discussed in detail in subsection 2.3.1.

fields—astrobiology and philosophy of science—from unstructured textual data. Their work, while illuminating, has important limitations—primarily in that it uses the shared interest in a topic as a proxy for the social connection. This approach overcomes important biases of studies that construct social networks based off of a citation network (since citations can be biased through various standards of the field), but nevertheless—inferrnig the existence of actual social relationships on this basis is highly fallible. Hence, it provides a very indirect proxy on the actual social practices in those fields.

While there are different ways the digital methods are increasingly used to center the scientific practice (e.g., through the study of "grey literature", such as peer review reports, see Hubert 2024; Miłkowski 2023a; Rorot and Miłkowski 2024), here I would like to explore a novel alternative, that motivates the work in this thesis.

While scientific theories are standardly regarded as the outcomes of scientific research and evaluated on the basis of their epistemic qualities (Bueno 2016), they also play a *practical* role in the scientific day-to-day by shaping the research design (e.g., H. I. Brown 1996) or restricting the possibility space for investigation (e.g., Love 2013; M. Wilson 2006). This perspective is advocated by Marcin Miłkowski in his recent proposal of a "cognitive metascience" programme (Miłkowski 2023a). Cognitive metascience views theorizing as a cognitive practice (Callebaut 2013; H. Chang 2017; Cheon and Machery 2016). Miłkowski views theories as perspectival "cognitive artifacts", i.e., entities used to "maintain, display, or operate upon information in order to serve a representational function and that affect human cognitive performance" (Norman 1991, p. 11), that extend beyond (e.g., generalize or abstract) the available data. Through this lens, scientific practice is a kind of human cognitive activity, and can be analysed within the framework of cognitive science. It can be viewed as an instance of distributed cognition (Afeltowicz and Wachowski 2015; Giere and Moffatt 2003; Hutchins 2006; Osbeck and Nersessian 2014; J. Zhang and Norman 1994), where the"cognitive labor is distributed among lab members, or among multiple collaborating teams" (Miłkowski 2022, p. 221). Crucially, the external form of the cognitive artfiacts plays a central role in those contexts, as it is what enables this division of cognitive labour: the sharing of theories depends on (and is shaped by) their expression—in the majority of the cases, linguistic and often textual (for visual representations of theories see e.g., Elkins 2001; Meynell 2013; Perini 2005; Sartori, forthcoming). It has been recognized also by the STS scholars, for instance in the Latour and Woolgar's (1979) inclusion of a detailed analysis of the literature produced at the Guillemin's lab in their study.

This understanding of theory is at odds with the received approach, which views theories primarily as a collection of sentences (in a logical language, commonly equated with laws—the syntactic view), a set of mathematical models (the semantic view), or a complex cluster of formal components from these two domains accompanied by nonformal elements, e.g., natural language statements (the pragmatic view; the distinction of the three main characterizations of theories follows Winther (2021)). Except for the pragmatic view (a relatively recent addition), which admits the impact of scientific practice on theories, these accounts limit the understanding of the theory to that of the outcome of scientific process. Instead, if Miłkowski's argument is on track, theories and their textual representations constitute a part of the scientific practice, and can be studied through their roles in research (a point also raised by Miłkowski (2023a)—this partially tracks the conclusions of French (2020), without submitting to the eliminativism he advances).

This is particularly visible in the contemporary cognitive science and psychology, where some claim that the replication crisis that has swept in the 2010s through psychology and other disciplines of social science (see Hughes 2018; Ioannidis 2005; Open Science Collaboration 2015) is (at least partially) a result of a theory crisis (Irvine 2021; Levenstein et al. 2023; Muthukrishna and Henrich 2019; Oberauer and Lewandowsky 2019; Smaldino 2017). The claim of a "theory crisis" is that the removal of explicit theoretical considerations from scientific practice significantly hinders our ability to achieve reliable results. For instance, several authors highlight that experimental work in psychology and neuroscience has been motivated mostly by exploratory experimentation (e.g., Feest 2012; which in fact can rely on theory as well, see Colaço 2018) in a vast possibility space, effectively limiting the statistical power of any possible experiment (Oberauer and Lewandowsky 2019).

For biology and cognitive science, it is commonplace to argue that these disciplines do not rely on theories in their developments (for biology the exception is the theory of natural selection; see the textbook treatment in Rosenberg and McShea (2012), ch. 2; see also Keller (2000); Ross (2025); for an alternative view see Sober (2018); for cognitive science, while some general theoretical frameworks are present (e.g., the information processing framework), the theories are usually limited to individual fields making up cognitive science, see Hardcastle (1996)). While they do presume a collection of mathematical and computational models, their limited generalizability and lack of systematic structure stops the majority of philosophers from viewing them as *theories* (even if they adopt the semantic view of theories).

This makes the cognitive view suggested by Miłkowski particularly relevant to account for the variety of theoretical activities involved in model-building and the synthesis of biological and cognitice science results. Importantly, the account of theories as cognitive artifacts highlights the importance of concepts associated with a particular theory, the framework they construe, and the way they shape the scientific process. In particular, as mentioned in chapter 1, I adopt a psychologically motivated notion of concepts as "the basic units of knowledge; [concepts] can be considered as the cognitive part of categories and consist of the information people associate with classes of objects and entities in the world" (Fini, Falcinelli, and Borghi 2025; this definition is motivated by the broader psychological perspective presented in e.g., Barsalou 1999; Borghi et al. 2017; Borghi et al. 2019; Borghi et al. 2023; Murphy 2004).

Considering the limited use of “theories” in life and mind research, I opt here for the notion of a “conceptual framework.” I understand it following Yosef Jabareen (2009, one of the very few philosophical treatments of this notion; see also Cheon and Machery 2016) as “a network […] of interlinked concepts that together provide a comprehensive understanding of a phenomenon” and “support one another, articulate their respective phenomena”, and introduce “ontological, epistemological, and methodological assumptions”.[2] Crucially, conceptual frameworks impact the identification, individuation and classification of phenomena by scientists (cf. Kant [1781] 1998). In psychology and cognitive science it is now commonly accepted that there are “top-down” effects in various domains of cognition in which our concepts, previous knowledge, or expectations shape what we perceive and experience (e.g., Clark 2013; Clark 2016b; Hohwy 2013; for arguments against this point consider Firestone and B. J. Scholl 2016). Hence, considering again the cognitive metascience view, it makes sense to assume that some such effects might be present in the scientific practice as well. This makes conceptual frameworks into a crucial part of the scientific practice.

While these considerations do not alleviate the worries regarding the reliability of textual outcomes of science (I come back to this point and offer a more detailed answer in section 2.3 below), their role as an external representation of the conceptual framework is quite unique. Crucially, in contemporary science, textual publications constitute a central avenue for the dissemination of scientific results and researchers’ concepts. While oftentimes the latter are included implicitly, they impact other members of the scientific community, who construct their own research based on the reported methodologies and results—hence, in relation to those frameworks (e.g., consider the treatment of the concept of “temperature” in H. Chang 2007). Further, in relation to scientists participating in peer review processes (of publications or grants) or sitting on hiring committees, these textual representations significantly shape the available opportunities and the broader theoretical landscape (e.g., the conceptual framework included implicitly in a grant proposal or in a writing sample submitted as part of a job application can impact the opportunities for carrying out further research within that framework—e.g., if the evaluating committee strongly considers it as an incorrect statement of the problem).

---

[2] Jabareen relies on the definition of “concept” offered by Gilles Deleuze and Félix Guattari (Deleuze et al. 2014), a notion expressed in a philosophical language quite alien to the theoretical background of the current work. Deleuze and Guattari understand concepts as having multiple components and being defined by them, and understood only in relation to them (Deleuze et al. 2014; following the treatment in Jabareen 2009), but “hav[ing] no *identity* but only a *becoming*” (D. W. Smith 2012, p. 62, emphasis original). Nevertheless, Jabareen’s notion of a conceptual framework doesn’t depend on that definition directly, and as such can be adapted to a more standard understanding of a concept, including the one adopted here. In the current context, the identity of concepts is understood through some version of a prototype account, in line with Wittgenstein’s “family resemblance” (Wittgenstein 1968), that Thomas Kuhn has partially adopted (see Barker, X. Chen, and Andersen 2003), or close to the notion of “patchwork concepts” (Haueis 2024), which I will adopt and cover in detail in chapter 3.

In this context, the textual representations of conceptual frameworks achieve a borderline status. They do serve as epistemic outcomes of the scientific process, but at the same time are an inherent part of the scientific practice—a dimension which I will call for the purpose of the current investigation *the conceptual practice of science*. As such, from a philosopher's perspective, scientific publications can be approached from multiple angles. Some philosophical problems—e.g., in naturalistic philosophy of mind—require treating them as epistemic outcomes (e.g., when we rely on the research on embodiment to analyse the questions about the relationship between the mind and body, to use a well-trodden example, e.g., Clark 1996; Varela, Thompson, and Rosch 1991). Other questions—e.g., in practice-oriented philosophy of neuroscience—motivate us to look at those publications as elements of conceptual practice that can be supplemented by historical and ethnological research about the parts of conceptual frameworks that are more difficult to recover from the texts (e.g., consider the emerging debate in philosophy of neuroscience about the conceptual and methodological commitments of the standard "neuron doctrine", which views neurons as primary functional units of the nervous system, and the novel "population doctrine" which shifts the interest to neural ensembles, see Gold and Stoljar 1999; Saxena and Cunningham 2019; Yuste, Cossart, and Yaksi 2024; the work of Nedah Nemati offers a good example of this focus on the conceptual practice, including as embedded in textual outputs, see Nemati 2024).

It is the aspiration of the current thesis to develop a methodological toolbox that can extract the conceptual practice from scientific publications using the tools of digital humanities, as this approach enables a uniquely broad perspective on those practices (in this way, it also extends the experimental perspective on scientific concepts, as reviewed in Cheon and Machery 2016). Such tools are not suited to provide any sort of automated conceptual analysis (see the points raised by Pence 2025), and need to be supplemented with standard methods of philosophy of science. Nevertheless, they can establish what Charles Pence (2025) calls a "conceptual cartography", "an understanding of the relationships between and features of scientific concepts as those concepts are expressed within scientific practice." Any view on conceptual practice based only on published textual evidence will necessarily be biased and incomplete, and should ultimately be complemented with the investigation of other sources of evidence—e.g., grey literature or through ethnographic interviews with practicioners of science. Nevertheless, the computational methods paired with large datasets available (published articles are uniquely covered in this regard) can begin this investigation, trading the biases inherent in scientific publications for the breadth of the scope they enable. And if we supplement this approach with qualitative close reading (see section 2.4), we have enough footing to begin the reconstruction of the conceptual practice of biology and cognitive science.

2.2 MAIN APPROACHES IN DIGITAL PHILOSOPHY OF SCIENCE

Broadly speaking, digital philosophy of science encompasses two general approaches to analysing scientific texts: network analysis, with citation network analyses playing a particularly prominent role, and textual analysis (see e.g., Lean, Rivelli, and Pence 2023). Both are suited to different questions. Network analysis uses textual data to infer network structure of entities, usually people or texts, mentioned in those texts. Citation is perhaps the most straightforward example: a citation network represents how individual texts and authors mention each other. This provides a proxy for understanding sociological and conceptual structures of a given field, which are becoming increasingly detailed, as methods are being developed to automatically infer the sentiment or function of a given reference (e.g., Budi and Yaniasih 2023; Ikram and Afzal 2019; Jiang and J. Chen 2023; Lu, Ding, and C. Zhang 2017; Nazir et al. 2020; Yousif et al. 2019). Oftentimes, however, network analysis may involve other types of data, such as collaboration or correspondence (for historical figures, whose letters are available) networks.

Some of the recent examples include studies conducted by the group consisting of Thomas Pradeu, Mahdi Khelfaoui, Yves Gingras, and Maël Lemoine (Khelfaoui et al. 2021; Pradeu et al. 2024). They used citation network data to study how philosophy of science interplays with actual science. This led them to identify a trend of "Philosophy *in* Science" (Pradeu et al. 2024), which highlights a particular way of doing philosophy of science that allows its practicioners to meaningfully address scientific problems, and, in the authors view, indicates a methodological continuity between science and philosophy of science (in a stronger sense than traditionally assumed in naturalistic philosophy of science). Another example comes from Charles Pence's work (2022a). In this paper, Pence uses articles and correspondence published in the journal *Nature* at the turn of the 20th century to identify key players of the debate in evolutionary theory over the possibility of generation of new species through the gradual accumulation of small changes, as postulated by Darwin's natural selection. Without the access to structured citation data (for historical reasons, as citation standards where not established in the analysed period), Pence uses a Named Entity Recognition algorithm (a well-studied problem in NLP with a host of out-of-the-box solutions available) to extract the list of proper names mentioned in articles, and constructs a network of mentions. Pence's study indicates that the debate was much less influential than historians usually assume, and provides a novel perspective on the impact of such theoretical crises on the development of the scientific fields, connecting to broader discussions in general philosophy of science. Finally, social networks can be also constructed in a more nuanced manner, an approach recently developed in the work of Christophe Malaterre and Francis Lareau (2024b; Malaterre and Lareau 2024a), who use topic modelling (more on that later) to develop a "topic profile" of an author, and use similarity of those profiles to infer a social network. This allows them to identify what they call "hidden communities of interest," subsets of authors who share common scientific interests, but whose connections cannot be readily picked out due to the granularity of citation data. They support their model by looking at the evolution of communities of interest in the history of philosophy of science - the case study highlights that the emergence of communities, as identified by their approach, coincides with well-known episodes in the history of the field.

Network analysis is a powerful approach that allows to bring to light important relationships - social, conceptual - between authors and texts. It is particularly well suited to study temporal patterns of how fields develop, connecting to an important body of research in philosophy of science. In that it differs significantly from the methodologies of textual analysis, which focus on a much more fine-grained study of the scientific publications.

Compared with network methods, which focus on relations between texts, textual analysis usually foregrounds individual texts, even if they serve primarily the purpose of identifying more general linguistic, conceptual, or social patterns. As individual texts are the units of analysis in this thesis, and textual analysis serves as the basis of the methodology implemented for the conducted studies, in the rest of this section I will discuss this approach in detail.

2.2.1 *The distributional assumption*

An early linguistic observation that has shaped the development of modelling semantics in large language corpora has been formulated by John Firth (1957, p. 11): "You shall know a word by the company it keeps!" Firth discusses the role that collocations, namely patterns of co-occurrence of words, play in linguistic research and specifically lexicography. As his views are associated with a Wittgensteinian perspective on language, underscoring the relation between use and meaning (Wittgenstein 1968), this claim indicates both a methodological recommendation regarding the study of semantics, and has a stronger sense: "It can be safely stated that part of the 'meaning' of *cows* can be indicated as such collocations as *They are milking the cows*, *Cows give milk*." (Firth 1957, p. 12). While the discussion of the metaphysics of word meaning postulated by this "distributional assumption" is beyond the scope of this work, the methodological implication of Firth's postulate is quite straightforward: word neighbourhoods are a crucial source of data that highlights not only important syntactic features of the word, but may also have some bearing on its semantic contents. This approach has been in fact dominant in the Natural Language Processing [NLP] research community's attempts to provide a quantified description of the meanings in natural language (in this context, the contemporary prominence of this approach began with the work of Lund and Burgess 1996; T. K. Landauer and Dumais 1997; see also M. Mitchell 2019, ch. 11; Lenci 2008; Lenci 2018), commonly known as distributional semantics (sometimes also vector space semantics, for reasons which will become clear soon).

In digital humanities, it has two main incarnations: “simple” co-occurrence analysis and the more (mathematically) complex set of methods involving vector semantics.

2.2.2 *Co-occurrence analysis*

Co-occurrence analysis takes up Firth suggestion to look for neighbourhood of concepts, but limits itself to the exploration of notable co-occurring word sequences, collocations, which are picked up mainly based on their frequencies. This approach, with particular focus on the qualitative exploration of such notable co-occurrences, was employed in one of the earliest digital studies in the field of philosophy of science. In a paper summarizing his doctoral dissertation, James Overton (2013), delved into the role of explanation in science. Initially, Overton compiled a list of theoretically relevant terms associated with various aims of science, as discussed in the philosophical literature. Subsequently, he conducted comprehensive searches for those terms within a corpus encompassing all articles published in the journal *Science* over the course of one year. This approach furnished him with empirical evidence concerning the significance of explanations within scientific practice, with terms related to “explain” appearing in nearly half of his sample, more frequently than words pertaining to “theory” and others. Following a thorough, qualitative examination of a random subsample, Overton identified five distinct categories of explanations offered by scientists. He then proceeded to investigate how existing philosophical perspectives fared in accounting for these various categories.

A more recent example of the usefulness of these methods in digital philosophy of science comes from a study conducted by Christophe Malaterre and Martin Léonard (2023; 2024). In the first study (2023), the authors have used a list of 6 philosophically central epistemic—“theory”, “model”, “mechanism”, “explanation”, “prediction”, and “understanding”—and explored their frequencies and co-occurring words in the corpus of approximately 74,000 articles from the BioMed central database.[3] These results enabled them to investigate relationships between these epistemic terms (and the terms they frequently co-occur with), which eventually leads them to conclude (among other, more detailed consideration) that “model” is by far the most frequent epistemic term referenced by scientists, significantly more so than “theory” or “mechanism”, which eventually leads them to conjecture that models are the dominant “epistemic objects” of science, and that science itself centers around “fitting models to data to make the best possible predictions” (Malaterre and Léonard 2023, p. 24). The second study builds on the methodology and results of the first one (Malaterre and Léonard 2024). They expand the list of the initial 6 terms to 211 (e.g., “hypothesis”, “experiment”), which includes their collocations identified in the previous study.

---

[3] Available at: https://www.biomedcentral.com/, currently published by Springer Nature. The corpus they used corresponded to the full contents of the database in 2015, after filtering out short articles or documents with incomplete metadata.

These are split into 61 "semantic fields" based on the authors' expert judgment of the relatedness of the terms. Again, for these terms the authors explored their frequencies and mutual co-occurrences at the "field" level, which allowed them to construct a correlation network and identify central terms (using standard algorithms for network analysis). Overall, the results confirmed previous conclusions, e.g., the dominance of "model" as an epistemic object or "prediction" as the aim of science. The study of correlations, allows them to expand to more detailed observations. For instance, they are able to pick out the "identification of functions and roles [as] one major epistemic target for research on molecular processes or entities, notably proteins and genes" (Malaterre and Léonard 2024, pp. 160-161). Overall, despite using the same basic measures as in the study conducted by Overton (2013), the application of a more complex methodology and a much larger dataset has allowed Malaterre and Léonard to achieve a significantly more detailed conclusions—even if their reliance on purely quantitative data ultimately restricts those conclusions to validating already known problems in philosophy of science (an important criticism towards the standard digital humanities approach, see Hammond 2017; I discuss this point in greater detail below, in subsection 2.3.1).

Co-occurrence analysis has also found application in the work of Mark Alfano (2019), albeit in a markedly distinct context. Focusing on Nietzsche's moral psychology, Alfano sought to move beyond the examination of seemingly unverified and superficial connections among the various terms Nietzsche utilized, which, in his assessment, had dominated the secondary literature on this subject. To achieve this objective, he embarked on a diachronic study of Nietzsche's body of work, systematically investigating the co-occurrences of terms throughout his diverse writings. This investigation led to the construction of a conceptual network representing Nietzsche's moral psychology and underscored the changes it had undergone over the course of his career. Alfano's approach unveiled that certain concepts which had been the primary focus of Nietzsche's readers, such as "resentment" and "will to power," were infrequent within his oeuvre and remained disconnected from the central network of concepts he employed, as indicated by the paucity of their co-occurrences in close textual proximity.

What is common to those distinct topics and methodologies is that digital tools (finding collocations in large datasets) are employed to unveil usage patterns of relevant concepts that would otherwise remain unnoticed during a close reading of a limited selection of texts. Furthermore, these tools demonstrate that even "simple" co-occurrence alone may illuminate the concepts' critical semantic attributes, which any comprehensive analysis should encompass. This significance becomes particularly evident in the application of this methodology within contemporary lexicography (Horák and Rambousek 2018), where the practice of lexicographers is enriched by these methods and extended, for instance, by incorporating information about the syntactic structure of co-occurrences within distributional models (see Kilgarriff et al. 2014).

2.2.3 *Vector semantics*

While looking directly at co-occurrences can be highly informative, the development of distributional semantics led to the refinement of mathematical tools beyond simple frequencies. As such, the dominant vector semantics approach offers a particular representation of meaning and a set of computational methods to infer those representations from empirical data (Lenci 2018, p. 152), most importantly from increasingly large textual corpora available contemporarily. The representation is a vector encoding the probabilities of co-occurrences of words across linguistic contexts calculated from the corpus, which is further processed using dimensionality reduction techniques.

To convey the idea, the general procedure is the following: for a target word, its collocates are determined within a particular context window. Depending on the purpose, often a set of "stop words", i.e., most common words which are believed to be semantically uninformative (e.g., articles, conjunctives, prepositions, etc.), is removed. Then, the frequencies of the collocates are merged into a vector which is the geometrical representation of the target word: the n-component vector defines the position of the word in an n-dimensional (semantic) space, with each dimension corresponding to a word in the vocabulary. This has obvious advantages: for one, using this method we can quantitatively compare the meanings of various words. The simplest example invokes the "meaning" of the plural form of a noun: $v(\text{"cat"}) - v(\text{"cats"}) \approx v(\text{"chair"}) - v(\text{"chairs"}) \approx v(\text{"idea"}) - v(\text{"ideas"})$ (where $v(\cdot)$ gives the distributional vector of the corresponding word) showing that the plural's contribution to the meaning is consistent across words. Second, this representation provides a quantitative way of judging semantic similarity between words. While the regions of the semantic space do not have anything special about them, the neighbourhood of words provides us with human-interpretable insights about their semantic properties. This approach has been one of the main drivers of the contemporary success of large language models like GPT (OpenAI 2023) or Bert (Devlin et al. 2019; Mickus et al. 2020).

While currently there are multiple available algorithms for calculating word embeddings and the dominant methods rely on pretrained large language models (Incitti, Urli, and Snidaro 2023), their initial development relied on simple word counts. The count-based algorithms[4] represent individual words using their indexes and occurrence counts in the vocabulary of the analysed corpus. While these methods are limited in many important aspects, they have introduced a numerical representation of word meaning which allowed for the application of artificial neural networks and eventually led to more complex approaches. One of the most prominent proposals is the Word2Vec algorithm (Mikolov et al. 2013), a shallow 2-level neural network, which relies on a novel, continuous version

---

[4] E.g., (Harris 1981), or the still used term frequency—inverse document frequency (TF—IDF; see Spärck Jones 1972; Salton, Fox, and Wu 1983) which is calculated as a product of the term frequency within a document and of an inverse function of the number of documents in which it occurs.

of either the bag-of-words (CBOW) or skip-gram architectures. Crucially, both architectures allow for calculating the vector representation based on both individual word and its immediate context: in the CBOW algorithm, the embeddings encode the impact of the word on relative probabilities of surrounding words (within pre-selected context window). In the skip-gram architecture, the model calculates an embedding that best allows to predict the nearby surrounding words from the represented word—in a sense, an inverse of the CBOW approach. Overall, as these can be inferred by a shallow neural network provided with a large dataset, the Word2Vec has provided much needed efficiency to vector semantics, cementing their role in contemporary natural language processing and linguistics.

Within digital humanities, this approach has been implemented along two main routes: in the form of topic modelling, and as what can be simply called distributional semantics. So far, only the first has seen larger uptake in digital philosophy of science.

Topic modelling

Topic modelling has been developed to achieve human-interpretable descriptions of large collections of discrete data, such as text corpora (Blei, Ng, and Jordan 2003), leveraging the distributional approach. The idea behind the approach is to represent a whole document (be it a sentence, a paragraph, or a larger text) as a vector of real numbers. The earliest approaches relied on count-based distributional methods, such as the TF—IDF scheme, to construct this representation. To overcome their limitations, more complex methods have been developed—most importantly, the algorithm of latent Dirichlet allocation (LDA; Blei, Ng, and Jordan 2003).

The LDA is a Bayesian generative probabilistic model for text collections which models individual texts as random mixtures of underlying sets of 'topics', where each topic is characterized by a distribution over individual words (Blei, Ng, and Jordan 2003). To put it more plainly, the idea behind LDA topic modelling is that a text can be described by the things that are mentioned. If we take the current chapter as an example, we can see that the main topics are the natural language processing algorithms and the methodologies of philosophy of science. The LDA makes an observation that each of those human-identifiable topics can be described by looking at the probabilities of individual words under the assumption that there is in fact a reference to that topic in the analysed text. It is indeed much more likely that words like "latent", "Dirichlet", or "allocation" are mentioned in a text that refers to natural language processing algorithms, than words like "resentment" or "will", which, in turn, have higher probabilities in reference to the topic of Nietzschean philosophy.

Hence, the LDA algorithm attempts to perform the inverse task of clustering words together dependent on their empirical distributions in individual texts in order to discover the 'topics' in those texts. The scare quotes appear at this stage, as there's a lot of skill—and risk—involved in the proper preprocessing of the data (e.g., "stop-listing", that is excluding semantically irrelevant frequent words), hyperparameter selection (most importantly, the number of 'topics' has to be decided in advance), etc. which is necessary to achieve some sort of semblance of the LDA 'topics' (which are simply lists of frequent words) to human-readable, meaningful topics (see e.g., the criticisms of some uses of topic modelling in Allen and Murdock 2022). Nevertheless, when used properly, the LDA offers a tool that provides unique insight into the contents of a large body of texts.

More recently, following the success of transformer-based large language models, novel algorithms which follow the same general idea have been developed, most importantly BERTopic (Grootendorst 2022). The BERTopic algorithm uses a selected, pretrained language model to calculate vector embeddings of individual documents within the corpus. This applies the methods discussed below, to achieve a more detailed, semantic representation of the text. For these embeddings, standard methods are used for dimensionality reduction and clustering. Crucially, in this context, the topics correspond to identified clusters in the reduced embedding space. However, their representation is constructed similarly to the standard LDA algorithm: for each cluster an individual bag-of-words representation of word frequency is created, and a class-based version of the TF—IDF algorithm is applied to identify words that differentiate individual clusters the most (though the Python implementation of BERTopic allows some more extensive control of the representation as well).

The main advantage of this novel algorithm is that identifying topics in the space of semantic representations overcomes some of the limitations associated with topic-interpretability (or more broadly, the correspondence between topic-modelling 'topics' and topics as understood by human readers, see Allen and Murdock 2022) in the standard LDA approach, at the same time outperforming the LDA model on standard benchmarks, without a significant increase in computational complexity (see Grootendorst 2022).

Topic modelling has been the most prominent area of use of distributional methods in digital philosophy of science. Some of the recent examples crucial for the design of the methodology used in this monograph include the studies of Jaimie Murdock and colleagues (Murdock et al. 2017; Murdock, Allen, and DeDeo 2017), as well as the research of Christophe Malaterre's group (e.g., Malaterre, Chartier, and Pulizzotto 2019; Malaterre, Chartier, and Lareau 2020; Malaterre, Pulizzotto, and Lareau 2020; Malaterre et al. 2021; Malaterre and Lareau 2022, 2023; Bertoldi et al. 2024).

In Murdock et al. (2017), the authors advance an approach they call a "drill-down" methodology. Taking as a case study the use of arguments invoking anthropomorphism in early days of comparative psychology, they use topic modelling to identify key passages in a large corpus. With their corpus consisting of the public domain books available through the HathiTrust Digital Library (see https://www.hathitrust.org/), their methodology is as follows:

1. Keyword search: the researchers begin with a keyword search using a Solr index provided by the HathiTrust Digital Library to identify books that are related to their area of interest and reduce corpus size to one that is feasible for topic modelling. This allowed them to narrow down to a corpus of 1315 publications (a corpus they term HT1315).
2. Book-level topic modelling: for the HT1315 subcorpus, they constructed probabilistic topic models at the book level using LDA, modelling each document as a distribution over topics, allowing for mixed-membership (i.e., a single document being classified as belonging to more than one topic). The topic model was then queried in an exploratory manner, until a set of terms—a topic query combining 'anthropomorphism', 'animal', and 'psychology'—was discovered that allowed to identify the three topics that were relevant to the research question. By calculating the distance between the vector representations of individual books' mixtures of topics and the three relevant topics, the authors were able to filter the HT1315 to a much smaller subcorpus of 86 documents (HT86) that bears more direct significance for the particular research question of the project.
3. Page-level topic modelling: for the HT86, Murdock and colleagues constructed topic models using LDA at page level. As they note, page level was selected due to limitations in available metadata (i.e., a paragraph- or section-level model would be preferred as carrying greater interpretability). After querying the topic model, they discovered one topic as directly bearing on their research question of anthropomorphism. Using this topic, they ranked the pages according to strength of association and then used the first 800 pages from that ranking to rank the whole volumes from their subcorpus according to the number of pages within this 800 top pages concerned with anthropomorphism. This left them with a final subcorpus of 6 publications (HT6; the cutoff point was chosen arbitrarily).
4. Argument extraction: using semi-formal discourse analysis approach, the authors manually annotated the arguments contained on the HT6 pages selected using the page-level model, to generate argument maps.
5. Sentence-level topic modelling: the authors selected one of the HT6 volumes, Margaret Washburn's *The Animal Mind* textbook (Washburn 1908), as it had the most pages related to the topic of interest, as indicated by the page-level LDA model from step 3. They performed a sentence-level LDA topic modelling of this publication and used an arbitrary sentence from the manually annotated arguments (step 4) to query that topic model in order to identify most similar sentences within the Washburn's textbook.

Along the practical outcomes of this work (e.g., the authors note that their method led to a detailed investigation of volumes which were not initially highly-ranked when performing the keyword search, which indicates the benefits of this methodology in identifying pieces of texts more directly relevant to the researchers' interests), the authors are able to identify several philosophically-interesting features of the analysed texts, including textual suggestion that Washburn's perspective on animal cognition was influenced by Kantian theory of cognition: in one of the passages identified in step 5, Washburn makes the distinction between spiders' ability to perceive objects as being at various distances, and their inability to *judge* those distances (Murdock et al. 2017, p. 14). Interestingly enough, while the paper is quite scarce when it comes to discussing results, presenting more as a methodological paper and a proof-of-concept study, the results were discussed in greater detail in a white paper published online (McAlister et al. 2014)[5], despite an explicit statement that "it was not a primary goal of the Digging by Debating project to produce new insights into the domain-specific content" (McAlister et al. 2014, p. 32). The authors make two philosophically-relevant observations: they note that analysed authors, regardless of some of their other theoretical commitments, acknowledge flexibility and variability of behaviour of individual animals, and underscore the role of studying the development of minds, which might be easier done in non-human animals.

In a different study, Murdock, Allen, and DeDeo (2017) investigate Darwin's reading journal to study his information foraging patterns. "Information foraging" refers to a particular model of how agents search for information, which applies the perspective of optimal foraging theory from behavioural ecology. This approach attempts to explain how individuals search for resources (e.g., honeybees for pollen) as an interplay of decisions to "explore", look for novel resource patches, and to "exploit", take advantage of previously found sources, while achieving a maximal result with minimal cost. In the case of information foraging, exploration refers to seeking novel ("surprising" in the information-theoretic sense) information, while exploitation refers to period of making use of already known information, reinforcing already held beliefs (i.e., seeking negative rates of "surprise"). In this work, topics identified at the LDA topic modelling stage are not investigated directly, but rather are treated quantitatively as probability distributions, and studied only on the level of the measure of diachronic change.

In this study, Murdock, Allen, and DeDeo (2017) extract the chronological list of books read by Charles Darwin from his reading notebooks. Using full-texts of those publications, available now in public domain, they construct (book-level) topic models and use the differences of distributions of topics in those volumes to measure local and global reading decisions made by Darwin. Using these methods, the authors are able to discover that Darwin's reading was indeed following an exploitation-exploration pattern, where the exploratory periods continually indicate the search for new information, to an unusual degree. The results are

---

[5] This is a project report, presenting the results of a study "Digging by Debating: Linking massive datasets to specific arguments", which was funded as part of the "Digging into Data" project, organized by research funding agencies from Canada, the Netherlands, the UK, and the US, see the note online: https://www.nsf.gov/news/news_summ.jsp?cntn_id=122726. The website of the "Digging by debating" project and the report are currently available only through the WaybackMachine service of Internet Archive: https://web.archive.org/web/20160821090213/http://diggingbydebating.org/. The available information suggests that the primary goal of the project was to develop and test tools for automating argument analysis.

validated as the automatically identified shifts correspond to major biographical events (e.g., the beginning of the work on *The Origins of Species* corresponds to a large increase in a text-to-past cumulative surprise, indicating a "global" (i.e., over the whole recorded reading history) exploration of novel information, see Figure 2.1). These results allow them to show how a form of cultural evolution occurs on the level of an individual. Crucially, they construct a comparison in which differences in the books *for Darwin* are juxtaposed with the differences in analysed volumes from broader, cultural perspective, by comparing the dynamics of those differences in Darwin's reading order and in the publication order of the books. This comparison (see Figure 2.2 and the discussion in Murdock, Allen, and DeDeo (2017), pp. 12-15) indicates that Darwin's pattern is much more exploratory, compared to the general ("culture's", identified with the publication order) pattern, which is almost exclusively focused on information exploitation, i.e., decrease in "surprise". The authors take this comparison to support a "remixing" hypothesis, which sees the role of the reader, Darwin in this case, as "sampling texts from different periods in such a way as to juxtapose thematically distant readings (…) bringing them into unexpected contact" (Murdock, Allen, and DeDeo 2017, p. 15).

More recently, the topic modetwenty-threet has been extensively applied in the works of Christophe Malaterre's group, particularly to the study of the history of philosophy of science. They conducted several studies which focus on the changes in topical profile of journals such as *Philosophy of Science* (Malaterre, Chartier, and Pulizzotto 2019; Malaterre, Chartier, and Lareau 2020) and *Biology and Philosophy* (Malaterre, Pulizzotto, and Lareau 2020) and in larger multilingual collections of journals historically central to the development of the discipline (Malaterre et al. 2021; Malaterre and Lareau 2022). These studies document the history of the discipline, highlighting the broader tendencies in its development, in line with standard, qualitative treatments. More recently they expanded the focus to history of science through work investigating the emergence of the field of astrobiology (Malaterre and Lareau 2023) and looking at the history of statistics (Bertoldi et al. 2024). In the latter case, the authors studied publication history of the journal *Biometrika*. Founded in 1901 by Francis Galton, Karl Pearson and W.F.R. Weldon as a venue for the discussion of statistical methodology, it is one of the first publications in the history of statistics, with a well-documented history (Cox 2001). Their approach is perhaps the epitome of contemporary topic modelling methodology:

1. Cleaning and preprocessing of the data: removal of editorial articles, table of contents, etc.; machine translation of non-english language publications; lemmatization; removal of extremely frequent and infrequent words, among others.
2. Topic model construction: explicit hyperparameter (number of topics) selection using both a quantitative measure (topic coherence), and qualitative, subjective assessment of resulting topics by one of the authors of the study; construction of the topic model.

Figure 2.1: Text-to-text (top) and text-to-past (bottom) cumulative surprise over the topic distributions of books read by Darwin. Negative slope indicates lower surprise and corresponds to exploitation, while positive slope indicates increase in surprise and corresponds to exploration. Shading identifies three epochs picked out by an unsupervised Bayesian model, and dashed lines mark important biographcal events. Figure and caption adapted from Murdock, Allen, and DeDeo (2017, Figure 1 on p. 13). Used on CC-BY License.

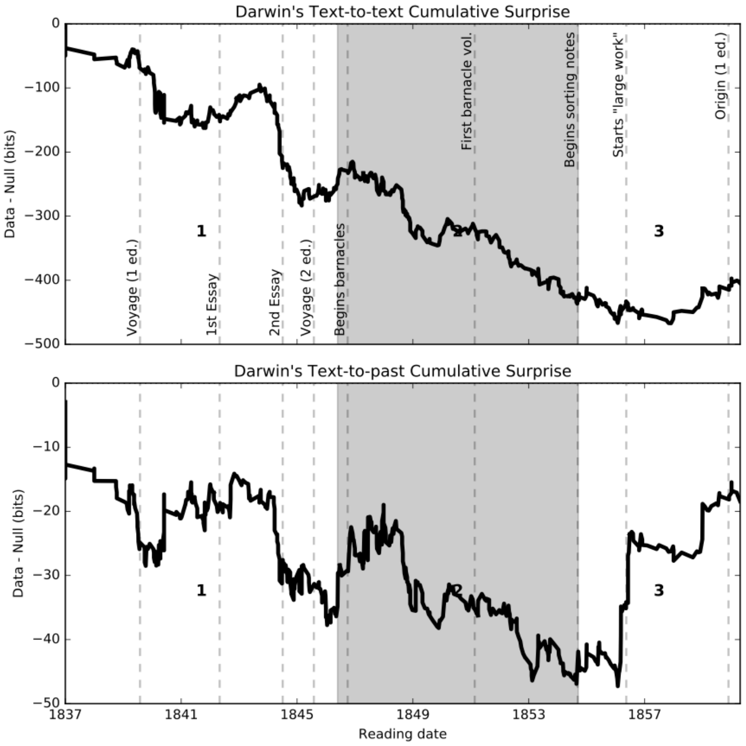
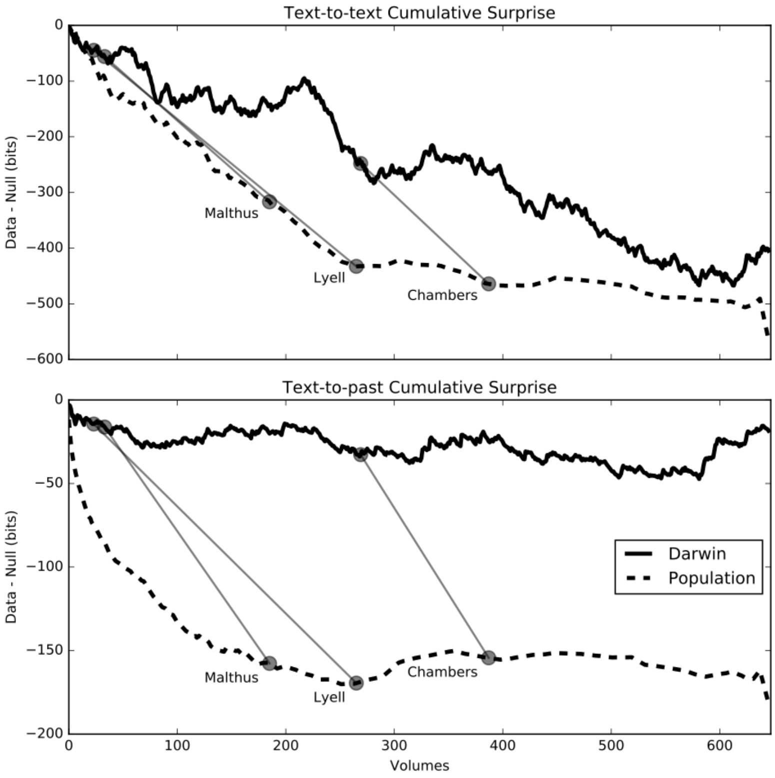


Figure 2.2: Comparison of text-to-text (top) and text-to-past (bottom) cumulative surprise over the topic-distributions of books read by Darwin, between Darwin's reading order (solid) and the publication order (dashed). Negative slope indicates lower surprise and corresponds to exploitation, while positive slope indicates increase in surprise and corresponds to exploration. Darwin's reading patterns indicate significantly higher exploration. Lines mark the positions of three biographically significant books: Charles Lyell's Principles of Geology (3rd ed., 1837; read in 1837), Thomas Malthus's An Essay on the Principle of Population (1803; read on October 3, 1838), and Robert Chambers's Vestiges of the Natural History of Creation (1844; read on November 20, 1844). Figure and caption adapted from Murdock, Allen, and DeDeo (2017, Figure 2 on p. 14). Used on CC-BY License.



3. Topic interpretation: the twenty-three resulting topics are manually assessed by looking at the top-10 words of each, and by expert review of the original articles in which the topic was most likely. This resulted in constructing short topic names, based on the top words, as well as more descriptive interpretations (with topics such as "Age-population" interpreted as "From vital statistics to mathematical statistics" or "Matrix-covariance" interpreted as "Methods for analysing multidimensional associations", see Bertoldi et al. 2024, Table 1). Further, for each article in the dataset, the topic distribution was calculated.
4. Topic clustering: the topic correlation was calculated and the resulting correlation coefficients were used to construct a network, that was further analysed using unsupervised clustering methods, to identify closely related topics (resulting in four topic cluster, corresponding to different stages in the historical developments of statistics (e.g., cluster A included topics such as "skull-measurement" or "correlation-mean", whereas cluster D included topics such as "model-effect"), presented in Bertoldi et al. 2024, Figure 4).
5. Diachronic modelling: using the individual article topic distributions and metadata about the publication date, the authors averaged over the topic probabilities of articles from discrete, five-year periods, to construct a diachronic topic profile of *Biometrika* (highlighting the changing prominence of topics associated with different clusters, e.g., with the probability of topics belonging to cluster A dropping to below 10% around 1960, concurrently with the rise of probability of topics from cluster B, see Bertoldi et al. 2024, Figure 5).
6. Author network: using methods mentioned above (see section 2.2), they constructed author network using topic distributions as measures of author distance, and investigated it further using clustering algorithms and by connecting the author and topic clusters.

While the final step is slightly unorthodox, steps 1-5 constitute the standard topic modelling methodology. Here, the results of this approach are validated by comparison with the mentioned existing historical accounts of both the development of the *Biometrika* journal, and statistical methodology more generally. Further, it provides additional support for those accounts that resulted from more focused close reading by indicating that those studies' focus on "major figures", such as Pearson, was indeed justified as a robust history of statistics, as the current historical story is not significantly revised by the computational study's inclusion of all the authors publishing in *Biometrika*. Furthermore, going beyond the history of the particular journal and discipline, the study provides an important insight about the social mechanisms, related to the structure of the community of authors, that underlie the historical development of scientific disciplines. In presenting their results, the authors exercise a large amount of caution. This somewhat limits their conclusions to confirming the pre-existing historical accounts—a source of criticisms for many computational studies in humanities, that I will discuss in detail below (see section 2.3).

Distributional semantics

While topic modelling has seen a wide application in all the domains that apply digital humanities methods, the main use case for other distributional method (distributionally-based collocations, or direct use of vector embeddings) has been in lexicography and in the study of concept change over history.

In the lexicographical case, corpus-based distributional methods are used to infer or disambiguate the meaning of the word, crucial in the case of polysemous words, to construct links between words based on their semantic similarity, and to identify occurrences of the word in corpus which can serve as good examples showing the word's usage pattern (see Fellbaum 2014; Horák and Rambousek 2018). One notable example is the SketchEngine tool and the Word Sketch algorithm it implements (Kilgarriff et al. 2014). Developed primarily for lexicographical purposes, Word Sketch offers a distinct approach to modelling word meaning by retaining syntactical information. While dominant approaches described in previous subsections represent a text as a "bag-of-words", this algorithm allows its users to distinguish between the same word serving different syntactical functions (e.g., "talk" as a verb or as a noun) and investigate collocations filtered by relevant syntactical relations, for example—look at notable (or typical) modifiers of a particular verb or noun. To establish the significance or typicality of collocations, the algorithm uses a purposefully designed measure of association between the words, known as LogDice (see Rychlý 2008; Kilgarriff et al. 2014). LogDice overcomes several issues of previously used measures, including the use of simple frequencies, as it enables e.g., cross-corpora score comparison, also for corpora of different sizes, as well as score interpretability (Rychlý 2008). For lexicographers, the benefits are quite straightforward—the Word Sketch provides them with corpus-based summary of the word behavior in language, allowing for detailed and empirically-based monolingual word definition, and cross-lingual meaning comparison. The tool has also a number of benefits for researchers in other fields of digital humanities, including digital philosophy of science (see Miłkowski 2022, 2023a; Rorot and Miłkowski 2024), as it provides quite straightforward qualitative information about the importance of word collocations, which in case of philosophically-burdened concepts can highlight relevance of individual terms, and enable a more detailed assessment of their meaning patterns, similarly to the methods described below.

One example of philosophical use of SketchEngine is our study of the concept of "scientific understanding" (see Rorot and Miłkowski 2024). In this work, we have used a relatively straightforward setup, basing on the Directory of Open Access Journal corpus available on SketchEngine, supplemented with our corpora of open peer reviews from the journal *eLife* (Miłkowski and Jasieński 2022), to identify common modifiers of the word "understanding" in the context of science. This allowed us to note the recurrent references to "predictive understanding". This escapes the existing philosophical theories which standardly connect understanding with explanations (e.g., Elgin 2007; Khalifa 2017; Potochnik 2017), rather than predictions (one exception is Dellsén 2016). Effectively, the use of distributional methods allows for the discovery of a novel, "distinct form of understanding that science can furnish" (Rorot and Miłkowski 2024, p. 271).

In digital humanities, distributional semantics have been used primarily for historical studies. This is possible, as these methods allow not only for evaluating syntactic and semantic use of words over a single corpus, but also for modelling the change of a particular concept between different subcorpora—whether they are organized diachronically, or by some other property. In this way the authors of two similar studies (Kim et al. 2014; Kulkarni et al. 2015) have investigated English corpora (including the Google Books Ngram Corpus, Michel et al. (2011)) to develop methods for deciding whether (Kulkarni et al. 2015) and when (Kim et al. 2014) does a shift in word meaning occur. These studies rely on the more commonplace vector embedding approach described in the context of topic modelling. The main methodological difficulty that these studies are met with concerns the ability to compare the word vectors between different semantic spaces, as these are created from individual corpora. The approximate methods used for calculating word embeddings can "produce vector spaces where words have the same nearest neighbours but not with the same coordinates" (Kulkarni et al. 2015, p. 4), since the coordinates are arbitrary. One possible solution to this problem is qualitative: investigating the membership of the neighbourhood of the studied word can highlight the semantic changes it has undergone over time. The other, quantitative solution, requires aligning the spaces to make the distance measures between words in different subcorpora interpretable.

This is where the two discussed studies differ significantly. Kim et al. (2014) do not acknowledge the issue of aligning the representational spaces explicitly and try to circumscribe it with the design of the word vector learning procedure. They begin with learning vector representations for the earliest subcorpus and then, for each subsequent corpus, they initialize their model with the representation learned from the previous one. This method is akin to iterative fine-tuning of the model, and in principle should ensure that the latent semantic spaces are approximately aligned, although there is no guarantee for the bounds on distances of the words which do not shift their meaning between subcorpora. Importantly, their particular use case provides theoretical foundation for this approach, as in any model of historical language change, the future states of language depend, at least partially, on its history.

Kulkarni et al. (2015), on the other hand, do discuss this issue, and propose a procedure for aligning the distinct semantic spaces of the subcorpora. Their approach has limitations—it assumes that the spaces are equivalent under a linear transformation and that the majority of words does not change their meanings between subcorpora—but provides a stronger guarantee that the spaces will in fact be aligned, as it boils down to inferring this assumed linear transformation with a regression model, while preserving the neighbourhood structure of the space (which follows from the second assumption).

With these methods, the authors of both papers are able to quantitatively characterize whether the change occurs and at what point. However, when it comes to classifying how the meaning changes between different subcorpora, they need to fall back to the qualitative methodology of looking at the nearest neighbours of the word of interest. With digital tools in hand, researchers can extract relevant information from huge sets of data and describe them quantitatively—building the interpretations on external evidence of the relevance of their selected sample. However, the task of word sense disambiguation, as technically this is classified, is difficult even for large transformer-based model which excel on other task (in a benchmark conducted by Basile et al. (2025); the top models did not achieve 80% accuracy). Given those limitations, ultimately, to answer the questions that humanities care about, one must return to the qualitative study of the available data.

2.3 CHALLENGES AND LIMITATIONS OF DIGITAL METHODS

While the digital methods offer unprecedented access to large swaths of data, they come with their own limitations and challenges. One source are the broader limitations of the methods of digital humanities that are still visible in their application to philosophy of science. These are further entwined in inherent limitations of the individual digital methods used. Further, there are more local problems related to the limitations of viewing science through the published literature (as I've already partially discussed in the context of conceptual practice, in section 2.1)—a set of problems that has recently been termed the "justificatory challenge" for digital philosophy of science. I discuss those two sets of challenges and limitations in turn.

2.3.1 *Interpreting results in digital humanities*

In the context of digital humanities, a significant restriction that has been identified in the literature is the preoccupation with methods, rather than results (most directly this criticism has been advanced in Kirsch 2014; see the discussion in Hammond 2017). These critics highlight that the work in digital humanities has failed to produce interesting results. Instead, the digital analysis of literature tends to reaffirm existing knowledge rather than propelling research in novel directions (Kirsch 2014). Adam Hammond (2017) acknowledges this fixation on validation while emphasizing that it is primarily focused on the validation of tools, thus neglecting the validation of results. In the context of the studies discussed here this is somewhat visible in the investigation into the history of statistics conducted by Malaterre's group (Bertoldi et al. 2024, see the discussion in Topic modelling). This study has led to a confirmation of an accepted view of the history of statistics, reaffirming the role of central figures. The use of digital methods undoubtedly offers an additional and independent confirmation of results achieved in a more standard, qualitative historical investigation, ensuring the robustness of those conclusions (e.g., regarding the central role of Pearson). Nevertheless, it is a somewhat disappointing conclusion for a study that offers an entirely new approach to the existing data.

Hammond terms this problem as the "double bind of validation" that has affected research in digital humanities: studies either utilize capable, validated tools but fail to yield groundbreaking results (as is the case in the study by Bertoldi and colleagues), merely restating what is already known or self-evident, or they produce astonishing and unexpected outcomes but struggle to validate and substantiate them. Hammond (2017, pp. 6-7) in discussing this gives an example of the research on a corpus of approximately 3000 British novels conducted by Heuser and Le-Khac (2012). These authors were looking for words increasing in usage in the course of the 19th century (replacing the words corresponding to "abstract values" like "envy" or "emotion", which decreased in usage after the end of the Romantic period) and found a group of words they called "hard seeds"—action verbs, body parts, colours, and numbers. While they expected the vocabulary closely related to the Victorian era (i.e., British imperialism or moral prudishness) they ultimately explained the surprising result by pointing to the move towards literary realism based on empirical description. However, their interpretation was underdetermined by the results, as later an alternative explanation was proposed by Underwood and Sellars (2012) which pointed to decreasing usage of words with Latinate roots (corresponding to the "abstract values" group) and increased usage of Anglo-Saxon words (corresponding to the "hard seed" group)—a broader trend in the British literary taste in the 1800s. In result, these remarkable findings may not advance the ongoing debates effectively, as they were not validated by the authors.

Established disciplines have a variety of methodological tools to ensure the validity of both applied methods and results produced by them. However, the relative novelty of digital humanities research introduces an additional burden onto the scholars pursuing this line of work. While validation of tools is necessary, at some point the research focus must shift to actually applying those tools to novel problems—in which case it is necessary to ensure that the results produced are not false positives and are interpretable. Digital methods can draw in this regard from the methods used by computational linguistics. In this context the validity of the results is partially ensured by using reproducible and replicable (see Miłkowski, Hensel, and Hohol 2018) methods, i.e., clear description of methods used (together with the relevant parameters) and the sharing of code and datasets used for analysis (see Wieling, Rawee, and Van Noord 2018). More importantly, however, computational linguistics often rely on a mix of human-supervised and unsupervised methods (e.g., J. Chang et al. 2009). This explicitly introduces human expert judgment e.g., for a tagging of a subset of data or a human evaluation of parts of the results. These then can be used to evaluate accuracy of the unsupervised component—following established approaches from machine learning, such as cross-validation techniques (see e.g., Goodfellow, Bengio, and Courville 2016, p. 121). These approaches are, however, feasible only for a subset of studies in digital humanities broadly, and digital philosophy of science more specifically. They are especially relevant for studies that aim to test specific hypotheses about trends in the published literature, and less so for more exploratory studies (see the discussion in Pence 2022b). As Pence (2022b, p. 251) mentions in his discussion of the "testing" versus "discovery" distinction in digital philosophy of science, the "spurious correlation and implicit assumptions present in the analysis of big data" offer a bigger hurdle to studies aimed at discovery. Indeed, the discovery is all the more difficult, as we may expect a lower base rate of the phenomenon of interest (e.g., when we're interested in emerging trends in scientific literature, these will be by default less represented than an established position, and hence run a higher risk of false negative results), and our approaches need to adjust accordingly.

These difficulties are further exacerbated by the limitations of particular methods. While all uses of computational methodologies will be highly sensible to the contents, quality, and size of the corpus used, individual methods run their own risks. Consider topic modelling, which serves as a major component of the digital toolbox applied in this thesis, and is perhaps the most prominent method in digital philosophy of science (especially the latent Dirichlet allocation algorithm). The standard methods of interpretation of the results have been subject of criticism for over a decade. For instance, in a famous paper, Chang and colleagues (2009) note that the topic modelling algorithms do not in themselves ensure semantic interpretability of 'topics'. Indeed, these authors then present results of a study with human subjects which concluded that better performance of the topic models on standard metrics might be in contradiction to topic meaningfulness: models with less interpretable topics performed better in their study. While methods for interpreting results have changed since—for instance, it is now commonplace to pair the automatically produced topic representation with qualititative evaluation of (some) documents assigned to the topic to establish the interpretation—difficulties persist. Allen and Murdock (2022) note that the problem is rooted in relying on the interpretation of topic models in the presentation of the results (compared with e.g., their own approach of treating them as probabilistic representations of documents and abstracting from possible semantics, see Murdock, Allen, and DeDeo (2017); discussed above). They highlight the practice of omitting some not-so-easily interpretable topics when reporting the modelling results, the limitations of relying on a limited subset of top words for interpretation (as the 10-20 top words carry relatively small percentage of the whole probability mass), or the risks of treating the models as a categorization algorithm, while in fact they produce mixed-membership classification (Allen and Murdock 2022, pp. 108-109). Instead, they propose to regard these models as "context" models, to establish the robustness of results across different possible hyperparameter choices (different numbers of topics), and an explicit inclusion of human interpretation practices.

Indeed, the inclusion of the human researcher as an explicit component of a digital methodology in humanities emerges in these discussions as the single best step in overcoming some of the challenges. In the methodology developed for the purpose of this thesis, I do indeed attempt that, by combining digital methods with standard, qualitative interpretation practice of philosophy of science—this large-scale design is discussed in the next section, section 2.4, and then followed by a more detailed outline of the study design in section 2.5. Before turning to this issue, however, I wish to address the second group of challenges to digital methods local to their application in philosophy of science.

2.3.2 *Justificatory challenges for digital philosophy of science*

The relationship between scientific publications and other elements of the scientific process is particularly relevant from the practice perspective, as discussed above in section 2.1. Nevertheless, it relates to several questions that need to be answered by any philosophical project that departs from individual publications (or even large corpora of those) and aims to arrive at generalizable conclusions. In an investigation like this, the relationship between the process of writing up scientific results into the form of conference presentations, articles, monographs, and eventually textbooks, and the actual scientific practice that leads to those results needs to be scrutinized. Publications certainly are a part of scientific practice (consider e.g., the view of the special role of textual artifacts emerging from the study by Latour and Woolgar 1979), but there is an ongoing debate of the exact role they play (Schickore 2008; Lean, Rivelli, and Pence 2023). Depending on a particular question under investigation, this tension may lead the researcher to consider moving away from studying published research outputs, and more towards other types of textual data sources, such as lab notebooks, study protocols, journals' peer reviews (e.g., Rorot and Miłkowski 2024) or even blog and social media posts. In the context of the current project, the particularly relevant question is how the conceptual practices in published scientific texts relate to the other avenues they may appear in (e.g., in formal and informal supervision relationships in a lab, which partially consist in reproducing the conceptual framework).

Lean and colleagues (2023) identify three central questions that need to be considered by every project aiming to derive philosophical knowledge from the (digital) study of scientific publications. The first question concerns the relation of publications to the broader scientific practice. The second concerns how we can derive generalizations about scientific literature from digital results. The third addresses the role of those generalizations for the philosophical inquiry. This last question has been in fact answered in the foregoing discussion regarding the interpretability of digital empirical evidence for humanities research (subsection 2.3.1): there are very few areas where philosophy of science can directly benefit from outcomes of computational analysis. Vast majority requires processing the quantitative results through a human interpreter—and only then these can serve any philosophical function (e.g., in conceptual analysis or engineering, as we discuss in Rorot and Miłkowski 2024). The remaining two questions, however, require some further consideration.

To begin with the first question of relation between publications and practice: Lean and colleagues begin by noting the common observation that publications offer only a distorted view on the inner workings of science. The authors propose a tentative taxonomy of the various ways this particular practice of scientists connects to their other undertakings. Focusing on the connection of publications to the context of justification (Reichenbach 1938), they introduce a historically grounded distinction between three main possible views:[6]

---

[6] Lean and colleagues note also a possible "fraud" view (Medawar 1996), the most radical and the most difficult to relate to the program of digital philosophy, as it holds that papers outright misconceive the scientific practices, presenting them as instances of inductive inferences: data is collected and generalizations are proposed that best explain data, whereas the actual studies are conducted to test for preconceived hypothesis with a deductive method. Medawar points to several important factors, such as the role of values in science (e.g., Longino 1990) or the relation between theory and experiment (e.g., Brandon 1994). These are indeed not as straightforward as

1) *the logical view*, which highlights that the "gap" between publications and practice is introduced purposefully by authors. In that view any departures from the more inexact and messy context of discovery mark the ways in which scientists believe standards of science are applicable to their research. And it is these standards that should be explicated by philosophers (e.g., Reichenbach 1938; Suppe 1998);
2) *the sociological view*, which focuses on how papers are designed to engage with a particular research community, in order to convince others to presented data and research programmes more broadly, to answer possible criticisms—within a preexisting social structure of science, which determines how the community is likely to respond to presented data (e.g., Knorr-Cetina 1981);
3) *the narrative view*, which centres the function of scientific articles as sense-making devices, constructed to relay (local and restricted) stories that place the research within a broader context and develop a particular structure that links the history of the presented research to the present and further extends into the future (e.g., Rouse 1990).

As Lean, Rivelli, and Pence (2023) highlight, each of those perspectives offers a distinctive approach for the interpretations of conclusions drawn from the investigation of published work, which in turn relate to the questions posed by a particular study. For studies concerned with conceptual issues within sciences, such as the current project, indeed the already discussed cognitive metascience perspective offers a beneficial approach (Miłkowski 2023a). The "cognitive view", so to speak, can be considered an alternative to the views listed by Lean and colleagues. To recap, the cognitive metascience perspective applies the category of "cognitive artifacts" (Norman 1991) to the analysis of theories and their (textual, visual, computational, etc.) representations (cf. Callebaut 2013; Miłkowski 2022)—but the category is intended to work broader, including all scientific outcomes. This connects the "narrative view" listed above, in that cognitive metascience underscores the role of textual artifacts as sense-making devices, with elements of the "sociological view", as it underscores the socially distributed character of doing science—and hence the social role textual publications play in the scientific process, though it downplays the persuasive characteristics Knorr-Cetina (1981) focused on.

The benefit of this framing vis-à-vis the justificatory challenge is that it allows for a well-founded pluralism about the functions of scientific papers. As vehicles of cognitive artifacts, publications scaffold a variety of practices of science, including conveying justifications, convincing other scientists, and offering narratives about the research and studied phenomena. The biases they introduce will likely be shaped by cognitive biases of their human authors, and as such

---

the models of early philosophy of science tended to present them. However, his conclusion that scientific papers are a "travesty of the nature of scientific thought" does not provide any tools for studying scientific publications.

will have a systematic form that lends itself to studying in an established way (cf. Wimsatt 2007). At the same time, in many cases these biases will be embedded and merged with the very artifacts, in particular in case of the types of artifacts that have privileged relationship to publications, such as theories or conceptual frameworks.

Here I propose to regard concepts (understood as basic units of knowledge, consisting of information associated with classes of entities in the world) and conceptual frameworks (networks of interlinked, mutually scaffolding concepts), inasmuch as they can be distinguished from any particular theory (while remaining theory- and value-laden), precisely as cognitive artifacts (in the sense of Norman 1991; Miłkowski 2022, 2023a) that are largely shaped by the forms of their textual representations in published outputs of the scientific research.

Note that concepts are in this respect significantly different from other artifacts of science. For example, for methodological toolboxes the biases existing in publications play different and limited roles (e.g., socially as factors in gatekeeping). They are determined in large part by various forms of "grey literature", such as lab handbooks, equipment user guides, various training materials, etc., or even more importantly conveyed in direct interactions in labs (aside of the various STS studies cited above, see also e.g., Nelson 2018, for the investigation of "epistemic by-products" or "tacit" knowledge; and Schickore 2020, for in-depth discussion of the history of the methods discourse).

Notably, the textual oputputs then act as cognitive artifacts themselves. An important benefit of this is that it underscores texts' active and constructive role in relation to the practices that have created them. Rather than being a simple passive representation, publications serve to reinforce particular values, highlight some aspects of scientific investigation, while hiding others, and in the process of doing so, directly construct scientific concepts. This observation parallels Karen Barad's criticisms of reflexive methodologies, which partially stem from the observation that reflexivity assumes "that practices of representing have no effect on the objects of investigation" (Barad 2007, p. 87).

This proposal builds on the views of Ian Hacking (particularly Hacking [1983] 2007), whose entity realism was strongly founded on the detailed analysis of microscopy and the scientific practices it scaffolds, an example that Barad also picks up (e.g., Barad 2007, pp. 50-53). The representations constructed with contemporary scientific microscopes hinge on the tool's ability to interfere and diffract the objects and processes they are intended to image, actively constructing their representations. Hacking extends the analogy to other areas of scientific research, providing an epistemic framework within which to consider this active role of scientific practice. Importantly, it motivates also his views on scientific realism. He proposes to view as real the entities that can be manipulated "in order to experiment on something else" (Hacking [1983] 2007, p. 263).

Barad follows in Hacking footsteps, extending the microscope analogy to the scanning tunneling microscope (invented around the time Hacking was publishing *Representing and Intervening*, and bringing the issues surrounding "seeing" with a microscope to an extreme). However, they claim that Hacking's focus on experimentation "leav[es] theory behind, ensnared in the trap of representationalism" (Barad 2007, p. 54). Instead, Barad argues that "experimenting and theorizing [as] dynamic practices that play a constitutive role in the production of objects and subjects and matter and meaning" (Barad 2007, p. 56). Putting metaphysical considerations aside, an important methodological corollary of this view of scientific practice is its diffractive nature. As an alternative to the more standard, especially in social sciences, reflexive approaches, Barad offers such diffractive methodologies which centre patterns of difference, eschew passive representationalism, and appreciate that "practices of knowing are specific material engagements that participate in (re)configuring the world" (Barad 2007, p. 91).

Viewed through the diffractive lens, publications both produce and are part of the phenomenon, serving a role not unlike a transducer in the example of ultrasonographic apparatus in an example analysed by Barad: imaging the phenomenon (a scientific concept) while at the same time bringing it to life (Barad 2007, p. 194-202; see also Barla 2023). Barad's analysis is equally applicable to science, as it is to the philosophical investigation of science, where we may strive to apply diffractive methodologies that are self-conscious about their double involvement in capturing and producing phenomena. Such methodologies render obsolete any preconceived subject—object (or publication—practice) distinction at the heart of the "justificatory gap" (in the sense of Lean, Rivelli, and Pence 2023). Within a diffractive approach to digital philosophy of science, publications may serve as "diffraction gratings", highlighting patterns of difference which show and embody the various agencies at play in theorizing. In this context the dynamic, dialectical emergence of those distinctions is itself open to questioning and epistemically accessible, as a central element of both the phenomena, and our investigation practices.

The answer to the second question, about the possibility of generalizing from studying publications, is more straightforward in the context of the current project. While I will engage in a generalizing interpretation of the results, all such claims are limited due to the considerations associated with the selection of appropriate corpora (see below in section 2.5) and the necessary inclusion of the qualitative close reading step (see section 2.4), which somewhat limits the scope of data investigated most exactly. As such, the generalization put forward is intended to cover a "dominant" or "mainstream" position, and does not preclude the existence of individual contradictory views in the literature—which the granularity of the applied methods simply overlooks. Nevertheless, it is important to note that in some domains digital methods can be paired with statistical reasoning approaches (either standard significance testing or Bayesian methods). In those cases, these can ensure the generalizability of the results—especially when questions such as prevalence of a particular conceptual framework are the main focus.

2.4 MULTI-LEVEL METHODOLOGIES OF DIGITAL STUDIES

The various applications of digital humanities—and the proposed solutions to the challenges the field and its methods face—underscore a crucial element of the corpus-based approach: while it can provide robust, empirically grounded data for the philosopher, it must be complemented by meticulous qualitative analysis and interpretative work (see Murdock et al. 2017). The prevalent methodological framework, spelled out by Murdock et al. (2017) and widely, though sometimes implicitly accepted, involves a twofold investigation. In the initial step, digital tools are employed for what is often referred to as "distant reading" (Moretti 2013). This step allows for the extraction of high-level patterns that are discernible only at the level of the broader knowledge system and accessible solely through the use of computational tools.

However, for these results to serve as valuable resources for philosophers, they must then undergo a second phase of interpretative close reading. This phase encompasses not only the results themselves, typically presented in the form of visualizations, but also an examination of the texts from which they were extracted. It involves identifying relevant subsamples of the corpus that are amenable to close reading (i.e., of a manageable size) and are chosen based on objective criteria (or at least at random), as opposed to arbitrary selection, which is frequently seen in conventional case studies within the philosophy of science (see section 2.1 above). The use of digital methods ensures the broader relevance of the selected subsamples within the context of the philosophical inquiries at hand. And the application of the "close reading" methods allows, in turn, a human validation of the results, a step that is seen as a desired practice in digital humanities (see Pichler and Reiter 2022), and has obvious epistemological benefits as a part of philosophical investigation.

Hence, the digital philosophy of science does not reject the extant methodologies, techniques, and approaches of naturalistic philosophy of science, but rather seeks to supplement these methods to overcome some of their limitations, especially those related to sample size. At the same time, digital tools introduce their own biases and restrictions which need to be carefully accounted for in the interpretation.

In this context, I propose to use this "multi-level" or mixed methods (R. B. Johnson, Onwuegbuzie, and Turner 2007) approach as the most promising way of capturing the broad conceptual practices embedded in scientific publications. The approach combines a broad, quantitative exploration with a more detailed and focused qualitative analysis of individual examples. In this way, it juxtaposes the emerging results and conclusions from each component of the analysis. Eventually, this enables me to overcome both the limited scope of standard philosophy of science methodology (see the discussion in section 2.1 above), and the restricted interpretability of computational methods (see section 2.3 above). Importantly, this approach explicitly admits the impacts of the human expert judgment on the digital tools, which are unavoidable (and necessary)—for instance, in data curation, keyword selection, or topic interpretation. Instead of viewing this as a source of unwanted unreliability (as qualitative methods are sometimes regarded), the inclusion of a *perspective* aligns with the broader view of scientific knowledge adopted in this thesis (see the discussion in chapter 1). Effectively, the researcher explicitly becomes a (part of) measurement device, engaging with the phenomena (traditionally conceived as the object of study, cf. Barad (2007); Hacking ([1983] 2007)). The emerging views on human knowing (De Jaegher 2021; De Jaegher and Di Paolo 2007), drawing from indigenous epistemologies (Kincheloe 2011; Massimi 2022, ch. 11), underscore this step as a crucial element of any practice of learning about the world.

2.5 DATA AND METHODS

The digital philosophy of science is still an emerging field, with its own set of methodologies adapted from the broader field of digital humanities, and the current thesis introduces several novel methods that are more suitable to the conceptual considerations central for the current project. The methodology developed for this purpose has been adapted from a variety of existing projects within digital humanities and digital philosophy of science, with a particular focus on enabling a “diffractive” investigation (in the sense introduced in subsection 2.3.2) of the concept of COMMUNICATION. In the remainder of this chapter, I outline the methodology of the studies conducted for this project—chapter 3 and 4 rely on results of one study, and chapter 5 on results of another, but both those studies apply the same exploratory schema.

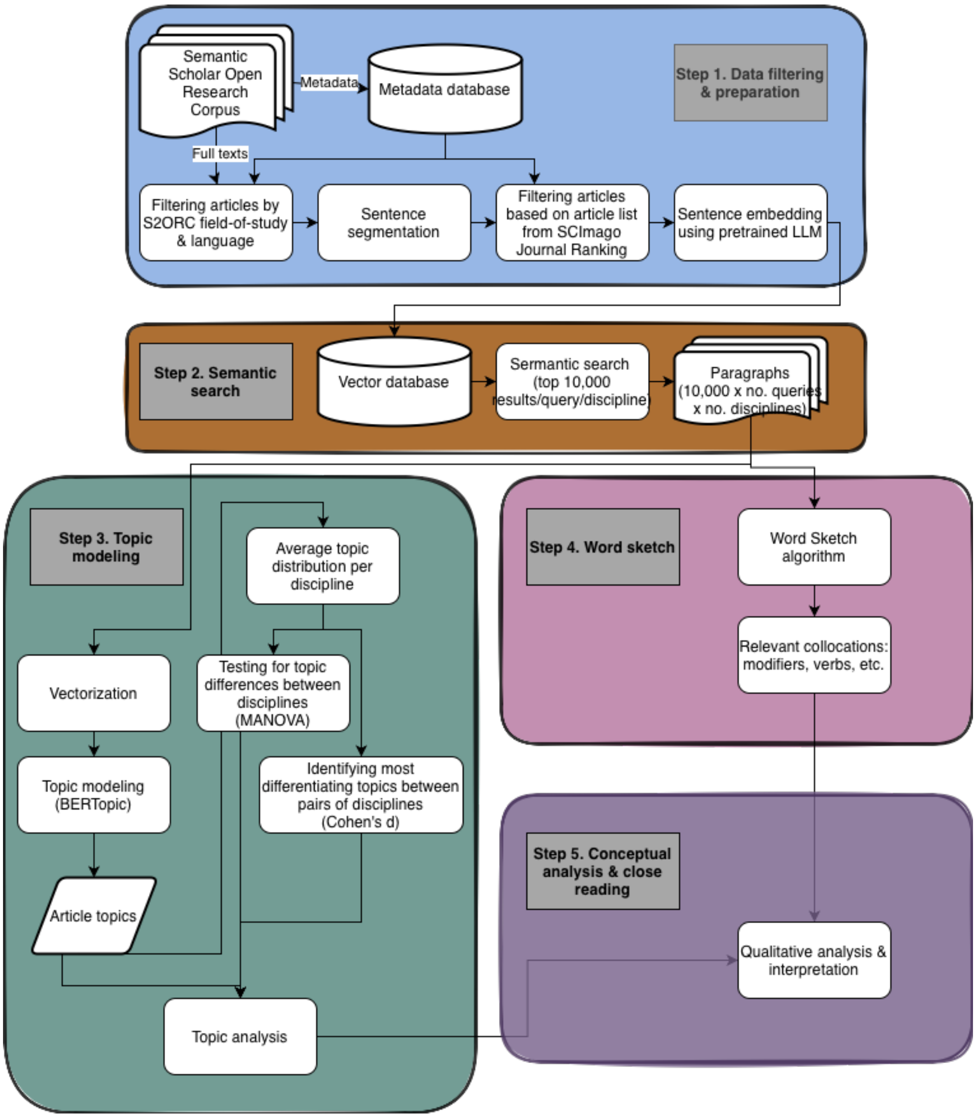
The research design is presented in Figure 2.3. The overall schema of the study was the following: after corpus preparation (Step 1. Data filtering & preparation), I have extracted passages (Step 2. Semantic search)—sentences and paragraphs—which mention communication (study 1—for chapter 3 and 4) and its semantics (study 2—for chapter 5), using semantic search methods (i.e., based on sentence-level vector representation of the corpus). This data was then processed in two ways: first, I have constructed paragraph-level topic models (Step 3. Topic modelling), to extract a high-level representation of the contexts which were identified with semantic search. Second, in parallel, the paragraph level data was also analysed with the Word Sketch algorithm (Step 4. Word Sketch), to identify the finer-grained semantics of relevant terms through their collocates. Finally, all three types of data—sentences and paragraphs produced by semantic search, topic models, and collocates—were qualitatively analysed and interpreted with a procedure inspired by standard conceptual analysis and close reading used in philosophy of science (Step 5. Conceptual analysis and close reading). In what follows, I discuss the methods used in each step in greater detail, highlighting some of the methodological considerations and decisions made. I begin by discussing the corpus selection.

2.5.1 *Corpus selection*

For the study, the Semantic Scholar Open Research Corpus (S2ORC) was selected (Kinney et al. 2025; Lo et al. 2020). Originally released in 2020, the corpus grew from initial 8.1 million structured full texts of open access papers to the size of over 12 million papers (as of 2024). It is updated and maintained weekly by the Allen Institute for Artificial Intelligence, and at the time of implementation of this project, constitutes the largest publicly-available unified collection of academic articles. The selection of this dataset was motivated by several factors:

1) corpus size and broad scope: in the case of the current study and its limited ambitions at generalizations (as discussed in subsection 2.3.2), a larger

Figure 2.3: Flowchart summarizing the research design. Individual steps are discussed in detail in the main text.



corpus increases the likelihood of including concepts that may be of philosophical interest and was prioritized over the representativeness of the corpus. Size plays a pivotal role here given that languages are not adequately characterized by ergodic and stationary statistics, implying that certain relevant terms may only appear in corpora infrequently, if at all (see Dębowski 2021).

2) corpus format and annotations: the articles are preprocessed from original PDF files into high-quality machine-readable text documents. The preprocessing strives to retain, where possible, information about the article structure, including stand-off annotations for abstracts, paragraph segmentations, headers, figures, tables, mathematical formulas, as well as citations and bibliographies. It is the largest challenge of dataset preparation to extract such information from the source data, hence the availability of a large dataset with such annotation is an important benefit for the project. Note that these operations have been conducted automatically, hence the data contains some noise, in the form of incorrectly recognized characters, especially in mathematical sections, wrong parsing of formatting, etc. In the context of the current project some of these limitations are overcome by the methods used (e.g., the semantic search based on vector embeddings is more robust with regard to typos than e.g., keyword search). Nevertheless, these issues increase the chance of false negatives (some data might be overlooked if it is effectively "unreadable" because of the noise).
3) corpus metadata: the dataset has rich metadata for each paper, including publication date and venue, disambiguated author identity, and, most importantly from the perspective of this project, papers have ascribed scientific disciplines (using both automatic ascriptions based on topic modelling of abstracts, as well as original information, where available).
4) copyright and ethical considerations: constructing a novel corpus can also be a legally challenging task, as individual publishers differ in policies regarding automated processing of text, even if a subscription is available or the use is otherwise enabled by existing provisions in intellectual properties laws. The S2ORC corpus is constructed with the explicit purpose of enabling computational studies, and its open-acces character allows for the reproducibility of conducted analyses.

An important limitation of the dataset is that it contains only open access papers. While this might further restrict the generalizability of the results presented here, there is some initial work suggesting that textual differences between open and closed access scientific journals are not major. For instance, Verspoor, Cohen, and Hunter (2009) have compared a sample of open access and traditional subscription articles for their syntactic structure and semantic contents (operationalized in terms of lexical terms used) and found no significant differences. Open access publishing is rapidly growing in importance. For instance, there is a major push from several large funding institutions for better access to scholarly literature (see the Plan S adopted across the European Union, European Science Foundation 2021), some studies indicate that Open Access articles constitute around half of the literature accessed by researchers, with OA journals achieving similar impact to subscription-based journals, as measured by citation indices (Björk and Solomon 2012; Piwowar et al. 2018). However, given the legal considerations mentioned above, this is perhaps the only possible approach to establishing a corpus of this size and scope, while retaining reproducibility.

Given the costs associated with OA publishing and a local character of existing incentives to publish OA articles[7] it is possible that there are systematic biases in place with regard to the groups of scientists who publish OA papers. These could in principle shift significantly the use of individual concepts or the scientists' epistemic commitments associated with them, especially given that the decision regarding investing in OA publishing itself is influenced by perceived disciplinary norms (see Moksness and Olsen (2017)). Hence, the corpus should not be regarded as representative of all published literature in life and cognitive sciences, and caution should be exercised when interpreting the study results.

2.5.2 *Methods*

To study the semantics of the concept of COMMUNICATION, as well as the function it plays in research across a variety of contexts, both topic modelling and distributional semantics were included within a broader multi-level, diffractive research design. The design is summarized in Figure 2.3. Here I will overview methodological decisions involved in the design of each step, reserving technical details for the online supplement. All the analysis were conducted in the Python programming language using purpose-built scripts, and for the semantic search Milvus vector database was used (Wang et al. 2021).[8] As noted above, two studies were conducted for the purpose of this project and both follow the same methodological schema, but involve additional queries and comparisons—these differences are discussed in each chapter, as necessary.

Step 1. Data filtering & preparation

The study uses a version of the S2ORC corpus published on August 6, 2024. The whole S2ORC corpus was filtered to select English language articles that are assigned the discipline of "Biology" or "Psychology" (see Table 2.1; cognitive science was not among the disciplinary labels used in S2ORC). This decision was based on the coarse-grained character of disciplinary assignments in S2ORC and

---

[7] The abovementioned Plan S is adopted mostly in European countries, and exclusively in countries of the Global North; the US follows funder-specific "public access policies", implemented both by National Institute of Health and National Science Foundation, two largest funding bodies; however in some Asian countries with large research communities, as in Korea or Japan, similar requirements are not implemented as of June 2025.

[8] Technical details regarding data processing and software used are published on OSF as an online supplement: https://doi.org/10.17605/osf.io/ax4jm. Code is available on GitHub (see the online supplement for details).

might have led to the exclusion of articles within traditional fields of cognitive science, for example artificial intelligence—which in S2ORC is likely to have been assigned the discipline of "Computer Science". The goal was to limit the noise (i.e., irrelevant uses of "communication", beyond the biological and cognitive context) in the dataset: for example, the inclusion of the whole discipline of "Computer Science" would result in including papers discussing e.g., computer network protocols, which make heavy use of the terms central to this study (COMMUNICATION, "signalling", "connection", etc.) but in a distinct, non-biological context—as uncovered in a preliminary investigation of the corpus. Similar limitations applied to the other disciplines standardly considered a part of the cognitive sciences, linguistics and philosophy (anthropology is not recognized as a separate field of study in the S2ORC), the risk of a high noise-to-signal ratio decided on their exclusion from the analysed dataset.

Table 2.1: Count of articles in dataset per field of study, as labelled within the S2ORC corpus.

Field of study	Number of articles
Biology (only)	2,173,186
Psychology (only)	556,641
Biology and Psychology	34,532
TOTAL:	2,764,359

The data was further narrowed down to the work published within selected subdisciplines (see Table 2.2) by filtering by the journal name, based on the Scimago Journal Ranking (SJR) lists assigning individual journals to more fine-grained subject categories (lists for the year 2023 were used).[9] Table 2.2 summarizes selected research areas and corresponding SJR categories. The selection of disciplines was further intended to filter out technical sciences which are outside the scope of the current study (e.g., agriculture or forestry). The division has been primarily motivated by the spatial scales of biological processes that the areas of biology are concerned with. As such it should be regarded as an operationalization of the concept of "scales" for the purpose of current research, rather than a claim regarding disciplinary distinctions within life and cognitive science.

The filtered subcorpus was segmented into individual sentences using a rule-based approach. For each sentence, for the purpose of semantic search, a vector representation was calculated using a pretrained, general-purpose large language model (LLM; Devlin et al. 2019; Reimers and Gurevych 2019; T. B. Brown et al. 2020), selected based on the standard benchmarks, bearing in mind com-

---

[9] The lists are available online at: https://www.scimagojr.com/journalrank.php?year=2023 (accessed October 8, 2024).

putational restrictions of the available hardware (see the details in the online supplement[10]).

Step 2. Semantic search

Due to the size of the filtered corpus, it was necessary to limit the analyses to relevant parts of articles—where BIOLOGICAL COMMUNICATION is mentioned. Since the distributional methods represent words with their immediate neighbourhood, a proper selection of fragments will effectively boost the signal-to-noise ratio and will have a very limited impact on the achieved representation of relevant terms. While historically the dominant approach was to use keyword searches, these have important limitations (see e.g., Hoeber et al. 2024). Most importantly, constructing even an inclusive list of terms related to COMMUNICATION to perform a broad search would include presuppositions regarding the semantics of the crucial concept, begging the question on the purpose of the current study, and introducing unwanted circularity.

For this reason, I have decided to replace keywords with semantic search (see e.g., Guha, McCool, and E. Miller 2003) which instead of matching character strings, looks for words or phrases with related meaning. For this purpose, the vector representation of sentence meaning calculated in the previous step was used. These encodings where then queried with different sets of queries for each study (discussed in detail in chapter 3 and 5). For each discipline (see Table 2.2) under investigation and each query, the search has selected top 10,000 best matching sentences from articles (without duplicating results in case a sentence was a top result for more than one query or more than one area of interest). The algorithm uses cosine similarity as the similarity score, a standard metric for analysis of similarity between vectors (the score is the cosine of the angle between the vectors in the representation space, with a maximum of 1 for identical vectors).

The final number of results depended on the number of queries used for each study and the amount of overlap between queries. For the study 1, reported in chapter 3, this was 84,110 sentences and corresponding paragraphs, and for the study 2 in chapter 5 this was 16,485 sentences and corresponding paragraphs (sample results are presented in the following chapters, and full results are available in the online supplement[11]). Using available paragraph annotations from S2ORC, for each sentence produced by semantic search, the paragraph which contains that sentence was retrieved from the full dataset for further analysis. To limit the noise, sentences shorter than 16 characters were omitted from the results (the number was selected arbitrarily, to limit the appearance of short headers and similar fragments containing 2-3 words). The choice of 10,000 most similar articles for each query and discipline pair is arbitrary. This criterion was chosen to ensure the comparability across queries and disciplines (for example,

---

[10] The online supplement is available on OSF: https://doi.org/10.17605/osf.io/ax4jm.

[11] Online supplement is available on OSF: https://doi.org/10.17605/osf.io/ax4jm.

the average similarity score for the results of each query become interpretable in this context) and to provide with sufficient amount of data for the further steps of analysis. The alternative would be to select a cut-off point for the similarity score. But considering the lack of direct interpretability of cosine similarity score values, I judged it to be a more arbitrary approach than the former.

Step 3. Topic modelling

To investigate the contexts in which references to COMMUNICATION are made, I construct a topic model of the selected paragraphs, using the BERTopic algorithm (Grootendorst 2022). As discussed in subsection 2.2.3, while the algorithm has similar motivation as standard LDA topic modelling and uses analogous method for representing topics, the identification of topics is done based on a vector representation of the documents (calculated with a pretrained general-purpose LLM). BERTopic performs dimensionality reduction (in the current work, the default UMAP algorithm was used, see McInnes, Healy, and Melville 2020) and clustering (in the current work, the default HDBSCAN algorithm was used, see Campello, Moulavi, and Sander 2013; Campello et al. 2015) on the vector representations, to identify topics. Then, it uses a modified version of the TF—IDF algorithm for constructing topic representations, though here additional, non-standard topic representations are used as well.

For both studies, different decisions regarding hyperparameters of the model were made, and these are reported in the following chapters, where individual models are introduced. BERTopic differs from LDA topic modelling in that it can both infer the number of topics in an unsupervised fashion, and reduce the topics to a smaller number in a supervised fashion. However, the unsupervised approach has a preference for larger topic models, which are less interpretable. For this reason, it was paired with a topic reduction to a chosen number of topics. In both cases, model evaluation was based on standard topic coherence metrics, C_v and U_{mass} (see the comparison and discussion in Röder, Both, and Hinneburg 2015), and a qualitative evaluation of the topics. For topic reduction, the goal was to achieve a granularity/coherence trade-off. Where it was relevant, the robustness of the conclusions was tested across different numbers of topics and different choices of hyperparameters. Importantly, BERTopic introduces an additional "outlier" topic which groups documents that did not fit in any of the clusters, and as such is not interpretable. To limit the number of such texts, additional steps were introduced. In that step, the algorithms attempts to assign to documents initiall classified as outliers their best-fitting topics. However, a number of outlier documents usually remains. These are normally omitted from the analysis.

To enhance interpretability of topics, the default representation based on most frequent words (top 20 words[12], using a standard list of stop words, as described in subsection 2.2.3) was accompanied by more specialized and nuanced representations that BERTopic enables: 1) top 20 most frequent nouns, modifiers and noun-modifiers pairs; 2) top 20 most frequent verbs, modifiers, and verb-modifiers pairs; 3) top 20 best keywords (single words and pairs of words) selected with the KeyBERT algorithm (Grootendorst 2020) and filtered with the Maximal Marginal Relevance technique to reduce redundancy of the keywords (both algorithms are included in the Python implementation of BERTopic). Based on these topic representations and the documents that were assigned a highest probability of the topic, I have manually created topic names that are used for analysis. The topics in study 1 were also analysed quantitatively across areas of interest using a recently developed method (Lawley et al. 2023): topic distributions of articles belonging to the areas of interest are statistically compared with one another using multivariate analysis of variance (MANOVA). Area of interest labels (e.g., "developmental biology", see Table 2.2) are taken as independent variables, while the topic distributions are treated as dependent variables. For statistically significant differences found by MANOVA, the effect size of the difference is calculated using partial η^2, which measures how much variance of the linear combination of the topics is explained by independent variables (Tabachnick and Fidell 2014). In parallel, Cohen's d (Diener 2010) is calculated to establish which topics differentiate the areas the most, to implement a diffractive methodology and focus the analysis on the patterns of differences. This quantitative information is used for the qualitative analysis, as the interpreted topics are compared with one another to establish theoretically relevant classes of contexts. This step is partially informed by the algorithmic topic clustering generated by BERTopic, but manual analysis is crucial. Further, the qualitatively identified differences between topic distributions are analysed in detail.

Step 4. Word Sketch

To investigate the more fine-grained semantics of relevant terms, the passages identified using semantic search where also analysed with the Word Sketch algorithm (Kilgarriff et al. 2014) to extract detailed information about collocation patterns of relevant terms. The relevant terms were selected separately for each study and are reported in chapter 3 and 5. The Word Sketch algorithm produced lists of top 20 collocates from different semantic categories. Of particular interest are the modifiers of those terms and nouns modified by them, however other syntactic relations are also taken into account as necessary.

---

[12] For some combinations of hyperparameters, the topic representations were limited to top 10 words. All the topic modelling results, as well as trained models are available in the online supplement on OSF: https://doi.org/10.17605/osf.io/ax4jm.

Step 5. Conceptual analysis and close reading

Finally, the crucial conclusions are drawn through conceptual analysis and close reading of documents and passages identified with computational methods. This methodology retains several important elements of the "drill-down" approach of Murdock et al. (2017), as discussed above in section 2.4. The digital methods provide information regarding the contexts of use of the key concepts (Step 2. Semantic search), which are modelled with topic modelling (Step 3. Topic modelling, following the suggestion in Allen and Murdock (2022)). The selection of the passages for close reading is largely (although not solely) determined by the computational results: I have focused on results with the highest similarity scores, varying across the different research areas included (the methodology was enhanced by pseudo-random sampling of the results).

The data used for close readings is threefold: most importantly, the analysis centres the relevant paragraphs and sentences from the corpus (identified in Step 2. Semantic search). However, this data is supplemented with direct quantitative outcomes of computational analyses and with a variety of visualizations generated from those analyses. This is perhaps the most distinctive feature of *digital* philosophy of science: the results of the study are represented in a visual form which enables better grasp of those results. This is done in a variety of established and more experimental ways (see Hinrichs, Forlini, and Moynihan 2019), but it is important to note at this point that these visualizations are algorithmically created and act to mediate the inference, which is ultimately grounded by data extracted directly from the corpus.

The close reading methodology aims primarily at delivering an account of COMMUNICATION and evaluating the coherence of the concept and its epistemological benefits. While this step derives from widely adopted approaches in philosophy, like naturalistically-oriented conceptual analysis (in the sense discussed by essays in Braddon-Mitchell and Nola 2009) and conceptual engineering (particularly by the procedures used to assess concepts, see Cappelen 2018; cf. Rorot and Miłkowski 2024), its goal is not to establish a *definition* of COMMUNICATION (as discussed in chapter 1), as this is not something that digital evidence can produce (see a recent argument in this line in Pence 2025). Importantly, at the close reading step, the empirical data is contrasted with extant philosophical accounts from the literature, as well as with a set of case studies. These serve a "diffractive" role (in the sense discussed in subsection 2.3.2), as they allow for extracting various implications of the emerging view of BIOLOGICAL COMMUNICATION. This approach allows a fine balance between the descriptive and normative goals of the current project (as outlined in chapter 1, and discussed in detail in chapter 6).

Table 2.2: Areas within life and cognitive sciences analysed in the study, with the corresponding Scimago Journal Ranking (SJR) labels included in each area. The total number of articles that belongs to at least one of the identified areas is 1,099,571. Note that SJR lists allow for assigning multiple disciplines for a journal, so the article counts do not sum to the total of analysed articles. On average articles have approximately 1.40 discipline assigned, with a median of 1 and a max of 6. Areas and labels are ordered alphabetically.

AREA LABEL	ARTICLE COUNT (inclusive)	SJR SUBJECT CATEGORIES INCLUDED
Animal Behaviour	62,628	Animal Science and Zoology Insect Science Small Animals
Developmental Biology	27,215	Developmental Biology
Ecology and Evolution	129,045	Ecological Modeling Ecology Ecology, Evolution, Behavior and Systematics
Generalist journals	67,967	Agricultural and Biological Sciences (miscellaneous)
Microbiology	267,334	Applied Microbiology and Biotechnology Cell Biology Microbiology Microbiology (medical)
Molecular Biology	587,402	Biochemistry Biochemistry, Genetics and Molecular Biology (miscellaneous) Genetics Genetics (clinical) Molecular Biology Molecular Medicine Structural Biology
Neuroscience	153,377	Behavioral Neuroscience Cellular and Molecular Neuroscience Cognitive Neuroscience Developmental Neuroscience Neuroscience (miscellaneous) Sensory Systems
Plant Science	93,934	Plant Science
Psychology	154,651	Developmental and Educational Psychology Experimental and Cognitive Psychology Neuropsychology and Physiological Psychology Psychiatry and Mental Health Psychology (miscellaneous) Social Psychology

3

WHAT IS COMMUNICATION?

The goal of this chapter is to analyse the concept of COMMUNICATION across the different scales in biology and cognitive sciences. As indicated in the Introduction, I will argue that the various uses of the concept preserve a core meaning, based on the empirical results from the corpus. To this end I will develop an argument by contradiction: using the notion of "patchwork concepts", one of the most widely used approaches to analysing polysemy in contemporary philosophy of science (see Haueis 2024), I will assume that COMMUNICATION has distinct senses across the different contexts. Indeed, the results of computational analyses of the corpus data indicate that there are many "patches" across which the concept is employed. If that is the case, then the definition-schema proposed in the Introduction (section 1.2) should be considered a general reasoning strategy, guiding the extension of the concept to novel patches, and providing pragmatic unity to the different uses. However, while the notion of patchwork concepts assumes that the patches are connected only by local semantic relationships, the account of communication emerging from the analysis does highlight a "core meaning" shared across *all* patches. Such a core meaning is taken to be an indicator that a concept—COMMUNICATION in this case—is not in fact a polysemous term. Incidentally, this emerging meaning aligns well with the proposed definition-schema.

Investigating the individual applications of the notion of COMMUNICATION brings to the fore the connections running through them. Based on the corpus data, I thus submit a novel account of BIOLOGICAL COMMUNICATION. Importantly, the shared core meaning of the concept allows it to play the same epistemic roles across different scales, primarily by identifying a shared organization of biological processes. To highlight the conceptual benefits of the proposed view, I analyse the use of the notion of "signalling" in the corpus, as a special case of biological communication. However, the view emerging from the analysis turns out to be too liberal. Hence, in the remainder of the chapter, I scrutinize the proposed account to identify what distinguishes communication from non-communicative causal processes and what distinguishes instances of biological communication from the broader category of "natural signs."

The chapter is structured as follows. In section 3.1, I introduce in detail the patchwork approach to scientific polysemy. In section 3.2, I introduce the analysis of the corpus data, following the methodology laid out in chapter 2 and discuss the results, adopting the framing of patchwork concepts to track the different uses of COMMUNICATION. Section 3.3 focuses on the shared elements of the different "patches" and argue that there is indeed a core meaning of BIOLOGICAL COMMUNICATION. Finally, in section 3.4, I adopt the emerging view to analyse how the term “signalling” is used in the corpus, and conclude in section 3.5 by considering the boundary cases challenging the proposed account, and effectively refining the view of COMMUNICATION.

3.1 PATCHWORK CONCEPTS

While often seen by philosophers as a hindrance to progress and clarity of scientific work, polysemy is a widespread feature of scientific terms, and does not seem to be a major roadblock in the research practice. While in philosophy of language there have been many attempts at capturing polysemy (to mention only Wittgenstein's notion of "family resemblance", Wittgenstein 1968; see also the review in Borg 2025), the more specialized domain of *scientific* concepts has received less attention. In this area, the notion of "patchwork concepts" (originally introduced by M. Wilson 2006) has become widely adopted, as it offers a well-defined, formalized way of capturing the different meanings of a term into individual "patches" of application.

Philipp Haueis, developing a generalised approach of "patchwork concepts", motivated by earlier uses, considers polysemous terms as being systematically employed "in different ways, each of which depends on the length scale [i.e., spatiotemporal granularity] of investigation, what technique is involved, the specific domain of application, and which property scientists target" (Haueis 2024, p. 742). As such they do not have any "core" nor "total meaning", beyond the meanings associated with individual "patches." The notion of a "patch" has been adapted metaphorically by Wilson to highlight the idea that the individual uses or senses are connected solely by local relations. These local relations may result from the overlaps of the domain or scale at which the two patches are used. Effectively, patches identify different classes of phenomena that are captured with the same term.The terms acquire novel meanings when a "generalised reasoning strategy" is extended to novel cases. Such strategies have the form of instructions which "tell researchers what operations they need to undertake to reach an epistemic goal" (Haueis 2024, p. 746), i.e., how to arrive at a measurement of a property associated with the concept.

One of the examples Haueis (2024, p. 749-750) discusses, drawing on his own previous work (Haueis 2021), is that of a CORTICAL COLUMN. In this case, the general reasoning strategy instructs researchers to look for vertical structures in the brain and then determine whether neurons within those structures have a similar function. However, the techniques for specifying the structure and function differ significantly across spatial scales of neural organization. CORTICAL COLUMNS have been originally analysed at the scale of "hypercolumns" using electrophysiological recordings, with the similarity of function of neurons within the column being approximated via uniformity of the recorded responses (see Hubel and Wiesel 1977). Later the concept has been expanded to the neural meso- and microscale. At the microscale, the columns are identified *ex vivo* using Golgi staining, and the functional organization is inferred from the dominance of vertical over horizontal connections between the neurons (see Mountcastle 1997). Each of those scale-dependent applications of the CORTICAL COLUMN forms an individual patch.

Haueis proposes to model the meaning of an individual patch P as $\mathsf{P}^{\mathsf{s}}\langle \mathsf{t}, \mathsf{d}(\theta)\rangle$. This formal definition includes four elements:

1. The scale s, at which the patch is identified;
2. The experimental or modelling technique, t, which operationalises the general reasoning strategy of the concept at the scale s;
3. The domain d, which specifies entities at the scale s to which the patch can be applied; and
4. The property θ, which is targeted by the technique t and is used to assign members of the class d to the extension (in the logical sense) of the concept.

For CORTICAL COLUMN at macro- and microscale, Haueis (2024) defines the patches as follows:

- Hypercolumn$^{2-3\mathrm{mm}}\langle$ tangential recordings, [V1, V2, MT](sequence regularity) $\rangle$
- Minicolumn$^{30-80\mu\mathrm{m}}\langle$ golgi staining, neocortex(vertical connections $>$ horizontal connections) $\rangle$

Importantly, Haueis (2024, p. 758-759) details the normative constraints that enable the systematic application of polysemous, patchwork concepts. He lists three conditions:

1) switching of techniques (t) when extending the concept to novel cases, to keep producing reliable results;
2) splitting the domain of application (d) into homogenous domains differentiated by the most informative and reliable property θ;
3) updating the "semantic picture" to capture which properties (θ) the different patches refer to, and why these properties "are significant to reach an epistemic goal" (most importantly, whether they actually do apply to the relevant domain).

These conditions are independent of one another, so that, e.g., techniques can be shared across patches, as long as they produce reliable results, patches can be constructed with overlapping domains, and properties from different patches can be combined to achieve broader epistemic goals via multiscale modelling or scale-bridging explanations—contributing together to creating the identity of a singular patchwork concept.

The analysis of COMMUNICATION through the notion of a patchwork concept requires identifying the various patches of application. With the relevant patches at hand, the general reasoning strategy connecting them can be specified and evaluated. As it will turn out, in line with what I have already suggested in the Introduction that the patches are associated with the various disciplines of life and cognitive sciences.

3.2 THE MANY USES OF COMMUNICATION

In this section, I introduce and discuss the results of the computational analysis. In subsection 3.2.1, I present the queries used for semantic search in the corpus (Step 2. Semantic search, see the full description of methodology in subsection 2.5.2) and review the motivation for the selection. These queries are based on definitions of communication selected from biological literature, almost exclusively in the context of animal behaviour. Along with the queries, I present a selection of the results intended to illustrate the various uses of communication found in the corpus.[1] In subsection 3.2.2, I introduce the topic model of the results from semantic search (Step 3. Topic modelling) and develop an argument that topics identified by the BERTopic algorithm are distinct patches (in the sense of Haueis 2024) of COMMUNICATION, a first step of the argument by contradiction put forward in this chapter. In subsection 3.2.3, I connect close reading of the results with the collocations found by the Word Sketch algorithm (Step 4. Word Sketch) to find the local relations between the patches. This is the second step of the argument by contradiction which is continued in section 3.3.

3.2.1 *Definitions of* COMMUNICATION

As discussed in the Introduction (see section 1.2), most philosophers focus on how the concept of BIOLOGICAL COMMUNICATION is used in animal communication research, rarely making references to phenomena at other scales at which the concept is applied. In fact, the debates surrounding gene signalling or cellular communication are more often couched in terms of information, representation, or at times symbols.

To briefly recap the discussion from the Introduction (see section 1.2), while there is a number of distinct definitions of biological communication, in the existing literature they have been sorted into two broad categories (see Scott-Phillips 2008; Frick, Bich, and A. Moreno 2019):

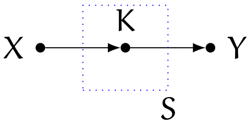
> **Informational**: COMMUNICATION is the process of information transfer from a sender to a receiver by means of a signal.
>
> **Influence**: COMMUNICATION consists in functional influence which a sender exerts on a receiver by means of a signal.

Quite clearly, there are several similarities between these two accounts, even if they pull theoreticians in different directions. These differences map onto the tentative, theory-neutral (with regard to the two definitions just listed) definition-schema of what kinds of processes qualify as communication that I have already anticipated in the Introduction:

---

[1] Full results are available as an online supplement: https://doi.org/10.17605/osf.io/ax4jm.

Figure 3.1: Diagram depicting the suggested organization of a communicative process.



> **Definition-schema:** COMMUNICATION is a process that involves a transfer or exchange of a certain mark K between distinguishable entities, X and Y, enabled by a structure S.

In the introduction, I have noted that this proposal is quite abstract, bordering on meaningless, and bracketed a detailed discussion until a later point. This point arrives just now. For the purposes of the argument in this chapter, I will assume that this definition-schema captures a general reasoning strategy underlying the ascriptions of communication to various processes across the scales of life (eventually, I hope to show that it actually captures more—a core meaning of the concept).

The account indicates three broad conditions that COMMUNICATION must jointly meet, and that are specified in a scale-dependent and technique-involving way across the different patches:

(1) transfer or exchange of a *certain mark* K,
(2) involvement of *distinguishable* entities, X and Y, and
(3) the presence of a *structure* S enabling such transfer or exchange.

The organization of the process is visualized in the Figure 3.1. Each of the conditions marks the parallels between the informational and influence views:

1) while the informational definition focuses on the transfer of (Shannon) information, the influence account cashes out the mark in terms of influence which, as we see in how "signals" are specified, refers to a certain perceivable behaviour, with vocalizations (e.g., alarm calls) being one example;
2) both definitions assume the existence of a "sender" and a "receiver" (in the influence context sometimes called the "producer" and the "consumer," see Millikan 1984). While most standard use cases refer to situations where these are clearly distinguished (i.e., different individuals), COMMUNICATION is at times applied also where such clear-cut distinctions are not suitable. One example is the reconceptualization of memory as "sending messages through time" (Godfrey-Smith 2012; see also the teleosemantic model of memory in Bielecka 2024). For this reason I decided to opt for the less restrictive notion of "distinguishable" entities (I discuss this in more detail in subsection 4.2.1);
3) the reference to structure in both contexts highlights that not just *any* mark will do. The informational approach relies here on a classical Shannonian schema, which views signal as an encoded message, transferred over an

established *communication channel* which constitutes (a part of) the structure (together with the encoding/decoding components). The influence approach posits a similar structure without referring to the notion of a channel, by stipulating that the signal must affect the behaviour of the receiver, evoking a response—i.e., that the receiver must be sensitive to the signal modality. It goes further, however, by requiring the signal-response pair to be functionally tied, hence the structure cannot be arbitrary. For example, in the adaptationist approach (as proposed by Scott-Phillips 2008), the structure includes the co-evolutionary trajectory of the signal response pair.

More broadly, these parallels between the views suggest that these two definitions are motivated by the suggested general reasoning strategy. If that is correct, the use of the term across different patches should be guided by these conditions, and they should motivate the choice of techniques used for identifying communicative processes. Further, they should be more strongly associated to the discussions related to COMMUNICATION than to mentions of picking out cues or some more directly causal relationships, which could potentially be conceived as competing conceptualizations of the kinds of processes that are captured by COMMUNICATION. Nevertheless, the definition-schema, to act as a general reasoning strategy and not a core meaning of the concept, should remain underspecified in the actual scientific uses. I will now turn to the presentation of empirical results, which—as we will see—fail the final requirement.

Importantly, the informational and influence-based definitions offer a good foundation for queries to use in the semantic search (Step 2. Semantic search, see subsection 2.5.2). The goal here is to develop an empirically grounded account of what communication is in biological and cognitive research. Querying the corpus for sentences that are semantically related to the definitional statements of the sort analysed here provides a way of identifying the sentences where the notion of communication is being directly referenced. Further, exploring paragraphs those sentences are embedded in offers a proxy to the context in which it is considered appropriate to describe a particular process in terms of COMMUNICATION.[2]

To avoid too narrow results, I have queried the corpus with multiple definitions of COMMUNICATION proposed in existing theoretical literature (presented in box 3.1), which mostly split along the two general approaches identified by Scott-Phillips (with important exceptions, as noted in box 3.1). The main source for the definitions has been the Historisches Wörterbuch der Biologie (HWB, Toepfer 2011) and the associated BioConcepts database (Toepfer 2024), a rare resource (mostly in German), which defines central concepts of biological research,

---

[2] Before deciding on this approach, I have explored the possibility of using more general, custom-made statements about communication to query the corpus. Qualitative analysis of those initial results indicated that using more precise sentences—such as actual textual examples of the definitions—offers more accurate results, i.e., less artifactual sentences are picked out in semantic search.

providing an overview of their history and examples of definitions from the literature. HWB and BioConcepts were designed as references for biologists and researchers working in related fields and provide a dictionary of central concepts used in biological research, based on their actual uses, and received positive reviews in this context (see Schmieder (2012)). While likely not exhaustive and focused only on major sources, HWB and BioConcepts are unique resources that provide a broad overview of the definitions of COMMUNICATION provided in the literature (mostly in the context of animal behaviour studies, as noted previously).

These definitions have been supplemented with definitions of "signalling" from the same sources, which are classified by the HWB under the category of communication (i.e., in the same entry), and apply analogous structure as definitions of "communication." Full list of used queries and their sources is presented in box 3.1. The search resulted in 84,110 sentences and corresponding paragraphs, balanced across the queries and disciplines. Representative results of the semantic search are provided in box 3.2.

Box 3.1: List of all textual examples of definitions of communication used to as queries for semantic search.

There was partial overlap between HWB (Toepfer 2011) and the associated BioConcepts (BC) database (Toepfer 2024), as well as between examples of uses of "communication" and "signal", and these sentences were not duplicated. For BioConcepts, only examples from the 20th century were selected, due to the temporal makeup of the corpus (see chapter 2).

Some sentences were slightly modified to reformulate them as directly definitional and to remove ellipses and unnecessary fragments that could impact the vector encoding of sentences. For example, the full version of the query 5_BC_DEF1_EX3 in the BC database is "In the most general sense, communication includes any stimulus arising from one animal and eliciting a response in another." I have removed the initial phrase to make the vector encoding more precise.

The IDs encode the database from which the query was drawn. Similarity score values are average cosine similarities between encoded vectors of the given query and the 10,000 results from the corpus produced by that query, and have been approximated to four decimal places.

1. ID: 0_INFO
   QUERY: Communication can be defined in terms of information transfer from a sender to a receiver by means of a signal.
   SIMILARITY SCORE: $\mu = 0.3733$, $\sigma = 0.1099$

2. ID: 1_ADAPT
   QUERY: Communication can be defined in terms of adaptive influence which a sender exerts on a receiver by means of a signal.
   SIMILARITY SCORE: $\mu = 0.4157$, $\sigma = 0.0789$

3. ID: 2_BC_DEF1
   QUERY: Communication is the transmission or exchange of information or knowledge between individual organisms.
   SIMILARITY SCORE: $\mu = 0.4175$, $\sigma = 0.1170$

4. ID: 3_BC_DEF1_EX1
   QUERY: Communication between animals involves the giving off by one individual of some chemical or physical signal, that, on being received by another, influences its behaviour.
   SIMILARITY SCORE: $\mu = 0.4431$, $\sigma = 0.0615$
5. ID: 4_BC_DEF1_EX2
   QUERY: Communication among animals involves the transmission of information or some other commodity from one participant to another.
   SIMILARITY SCORE: $\mu = 0.4653$, $\sigma = 0.0808$
6. ID: 5_BC_DEF1_EX3
   QUERY: Communication includes any stimulus arising from one animal and eliciting a response in another.
   SIMILARITY SCORE: $\mu = 0.4679$, $\sigma = 0.0794$
7. ID: 6_BC_DEF1_EX4
   QUERY: Communication is the phenomenon of one organism producing a signal that, when responded to by another organism, confers some advantage (or the statistical probability of it) to the signaler or his group.
   SIMILARITY SCORE: $\mu = 0.4311$, $\sigma = 0.1060$
8. ID: 7_BC_DEF1_EX5
   QUERY: Biological communication is the action on the part of one organism (or cell) that alters the probability pattern of behavior in another organism (or cell) in a fashion adaptive to either one or both of the participants.
   SIMILARITY SCORE: $\mu = 0.4340$, $\sigma = 0.0620$
9. ID: 8_BC_DEF1_EX6
   QUERY: The behavior of communicating is behavior that enables the sharing of information between interacting individuals as they respond to each other.
   SIMILARITY SCORE: $\mu = 0.4717$, $\sigma = 0.1212$
10. ID: 9_BC_DEF1_EX7
    QUERY: Communication is said to occur when an animal, the actor, does something which appears to be the result of selection to influence the sense organs of another animal, the reactor, so that the reactor‘s behavior changes to the benefit of the actor.
    SIMILARITY SCORE: $\mu = 0.4127$, $\sigma = 0.0655$
11. ID: 10_BC_DEF1_EX8
    QUERY: The transmission of a signal or signals between two or more organisms where selection has favoured both the production and reception of the signal(s).
    SIMILARITY SCORE: $\mu = 0.4336$, $\sigma = 0.0541$
12. ID: 11_BC_DEF1_EX9
    QUERY: Communication is any action of one organism that modifies the behaviour pattern of another organism; the passage of information.
    SIMILARITY SCORE: $\mu = 0.4325$, $\sigma = 0.1208$
13. ID: 12_BC_DEF1_EX10
    QUERY: Communication: The completion of corresponding signals and responses. Signal: Any act or structure that (i) affects the behaviour of other organisms; (ii) evolved because of those effects; and (iii) which is effective because the effect (the response) has evolved to be affected by the act or structure.
    SIMILARITY SCORE: $\mu = 0.4591$, $\sigma = 0.0887$

14. ID: 13_BC_DEF2
QUERY: Communication is the physical connection between the parts of a body.
SIMILARITY SCORE: $\mu = 0.4037, \sigma = 0.1159$

15. ID: 14_BC_DEF2_EX1
QUERY: Highly integrated systems in which a communication system exists between the different molecules; This means that the molecules are informed at every moment of what is going on around them and that they work not in one way but in a highly coordinated manner.
SIMILARITY SCORE: $\mu = 0.4250, \sigma = 0.0363$

16. ID: 15_HWB_EX1
QUERY: Communication necessitates the existence of a code shared between two or more individuals whose use is mutually beneficial to its possessors, i.e. increases fitness.
SIMILARITY SCORE: $\mu = 0.3917, \sigma = 0.0309$

17. ID: 16_HWB_EX2
QUERY: Communication is the transmission of information from one animal to another.
SIMILARITY SCORE: $\mu = 0.4175, \sigma = 0.0883$

18. ID: 17_HWB_EX3
QUERY: Communication is the process in which actors use specially designed [i.e. designed by natural selection] signals or displays to modify the behaviour of reactors.
SIMILARITY SCORE: $\mu = 0.4138, \sigma = 0.1061$

19. ID: 18_HWB_EX4
QUERY: Communication involves carrying informational content, which can be manipulated by the sender and differentially acted on by the perceiver.
SIMILARITY SCORE: $\mu = 0.4566, \sigma = 0.1083$

20. ID: 19_BC_SIG_EX1
QUERY: The signal usually involves little expenditure of energy by the sender, and has a large positive or negative effect on the receiver's energy expenditure.
SIMILARITY SCORE: $\mu = 0.4378, \sigma = 0.0676$

21. ID: 20_BC_SIG_EX2
QUERY: Signal is any behaviour that conveys information from one individual to another.
SIMILARITY SCORE: $\mu = 0.4785, \sigma = 0.0747$

22. ID: 21_BC_SIG_EX3
QUERY: A signal is a trait that is selected to manipulate an addressee to the advantage of the sender of this signal.
SIMILARITY SCORE: $\mu = 0.4242, \sigma = 0.0688$

23. ID: 22_BC_SIG_EX4
QUERY: Signals are traits whose effect on their bearer's basic fitness is non-positive, and whose positive specific effect on their bearer's (signaler's) fitness stems from the fact that they change information held by other individuals (recipients). The recipient's changed information eventually results in corresponding change in their behavior.
SIMILARITY SCORE: $\mu = 0.44718, \sigma = 0.0505$

24. ID: 23_BC_SIG_EX5
QUERY: We define a 'signal' as any act or structure which alters the behaviour of other organisms, which evolved because of that effect, and which is effective because the receiver's response has evolved.
SIMILARITY SCORE: $\mu = 0.4335, \sigma = 0.0792$

25. ID: 24_BC_SIG_EX6
    Query: Animal signalling is the use of specialized, species-typical morphology or behavior to influence the current or future behavior of another individual.
    Similarity Score: $\mu = 0.4408$, $\sigma = 0.0574$

26. ID: 25_HWB_SIG_EX1
    Query: Signals are behavioural, physiological, or morphological characteristics fashioned or maintained by natural selection because they convey information to other organisms.
    Similarity Score: $\mu = 0.4743$, $\sigma = 0.0538$

27. ID: 26_HWB_SIG_EX2
    Query: Signal is any behavior that conveys information from one individual to another, regardless of whether it serves other functions as well.
    Similarity Score: $\mu = 0.4599$, $\sigma = 0.0761$

28. ID: 27_HWB_SIG_EX3
    Query: Signals are externally visible features of animals—'designed by natural selection'—to influence the behaviour of other animals, a means by which one animal (the 'actor') exploits another animal's (the 'reactor's') muscle power.
    Similarity Score: $\mu = 0.4572$, $\sigma = 0.0541$

29. ID: 28_HWB_SIG_EX4
    Query: Signal is an action or structure that increases the fitness of an individual by altering the behaviour of other organisms detecting it, and that has characteristics that have evolved because they have that effect.
    Similarity Score: $\mu = 0.4328$, $\sigma = 0.0603$

Box 3.2: Manually selected representative results for queries from box 3.1.

Different queries produced overlapping results, with different cosine similarities. To select the examples, I have sorted the results by the highest cosine similarity (across all queries) and the query corresponding to that score is indicated below. This box is intended to illustrate results of the semantic search to the reader, all the analyses reported in the chapter were conducted using the full data.

As summarized in the Table 3.1, some queries produced on average better (more similar) results, hence the selection here is skewed towards the queries with higher similarity scores. As the vector encodings have likely impacted this bias, it is impossible to interpret.

The examples were picked qualitatively by the author, to show the diversity of the results from the corpus. The manual selection was limited to top sentences. However, I have attempted to show examples from all areas of research included in the corpus. For each area, I show results produced by different queries. Hence, in the selection process I have included also some examples with lower similarity scores (full list with scores can be found in the online supplement). The reference includes the CorpusID from S2ORC and the sentence number—I use this convention throughout the thesis when referring to examples from the corpus analysis, to distinguish them from other citations.

1. Query ID: 5
   Similarity score: 0.8085
   Research area: Animal Behavior
   Result: As communication usually occurs in a network of several animals in signalling and receiving range of each other [24], the emotional state of a caller may influence the behaviour of several individuals, addressees in direct interactions and bystanders alike.

Reference: 5566073, 24

2. Query ID: 20
   Similarity score: 0.7523
   Research area: Animal Behavior
   Result: Definitions of communication vary, particularly in terms of intentionality and information (Rendall, Owren, & Ryan, 2009). A signaler may intend on influencing a receiver, or consequences of signaling may be inadvertent (Seyfarth & Cheny, 2003). For our purposes, communication is simply a social event that includes a signaler (forager honeybee performing waggle dances) and a receiver (dance follower) behaviourally responding to a signal.
   Reference: 53307193, 11
3. Query ID: 24
   Similarity score: 0.7174
   Research area: Animal Behavior
   Result: Animals may communicate not only about internal states but also about external referents, eliciting also the appropriate reaction from the receiver in the absence of the stimulus triggering the signal in the sender.
   Reference: 205330370, 20
4. Query ID: 20
   Similarity score: 0.7130
   Research area: Developmental Biology
   Result: AVPR1A signaling has also been shown to be involved in human social and emotional behavior as demonstrated in a study using the AVPR1A antagonist SRX246 [9].
   Reference: 12622276, 17
5. Query ID: 20
   Similarity score: 0.6442
   Research area: Developmental Biology
   Result: In sum, RA signaling is a key regulator of anteroposterior patterning orchestrated by Hox codes in at least general ectoderm, central nervous system and endoderm of chordates.
   Reference: 10613111, 174
6. Query ID: 6
   Similarity score: 0.6312
   Research area: Developmental Biology
   Result: Intercellular communication is a key biological process that enables cells to coordinate their responses spatially and temporally to physiological changes.
   Reference: 229344625, 128
7. Query ID: 16
   Similarity score: 0.7800
   Research area: Ecology and Evolution
   Result: Whether or not animal communication should be conceptualised as information transfer has long been a contentious issue (Dawkins and Krebs 1978;Rendall et al. 2009).
   Reference: 148569416, 11
8. Query ID: 8
   Similarity score: 0.7708
   Research area: Ecology and Evolution

Result: Through cooperation and competition, people realize this communication and promote the establishment and harmonious development of interpersonal relationships [25].
Reference: 259260855, 160

9. Query ID: 4
   Similarity score: 0.7613
   Research area: Ecology and Evolution
   Result: Animal communication includes a vast array of signalling systems, ranging from the warning colouration of noxious insects to the complexity of human language.
   Reference: 18825221, 8
10. Query ID: 11
   Similarity score: 0.6943
   Research area: Microbiology
   Result: Based on a previously reported protocol (Murata et al., 2017; Li et al., 2021), a communication box system was used for inducing psychological stress.
   Reference: 258487683, 41
11. Query ID: 18
   Similarity score: 0.6849
   Research area: Microbiology
   Result: Specifically, transcallosal information exchange is thought to ensure a good balance between excitation and inhibition of contralateral brain activation during bimanual performance to suppress inadequate mirror movements [3,12,19].
   Reference: 251734952, 131
12. Query ID: 23
   Similarity score: 0.6484
   Research area: Microbiology
   Result: But why should there be cell non-autonomous (diffusable) signals as well?
   Reference: 262595822, 114
13. Query ID: 11
   Similarity score: 0.7312
   Research area: Molecular Biology
   Result: Thus, the present study suggests that the concept of communication channel modulation between the perceptual decision-making system and the motor system may be the key to understanding the mechanisms underlying the link between decision-related activity and decision-unrelated motor processes.
   Reference: 261394193, 182
14. Query ID: 27
   Similarity score: 0.6193
   Research area: Molecular Biology
   Result: Crosstalk between pairs of signaling pathways has been identified in healthy organismal development (1).
   Reference: 6585223, 13
15. Query ID: 14
   Similarity score: 0.5740
   Research area: Molecular Biology
   Result: At the molecular level, this has led to the emergence of multiple signaling systems perceived by both interacting parties.
   Reference: 264953331, 9

16. Query ID: 5
    Similarity score: 0.8103
    Research area: Neuroscience
    Result: Understanding the mechanisms for human communication, and how it may go awry, requires the use of model animals that naturally exhibit at least a subset of similar communicative behaviors.
    Reference: 5880834, 304
17. Query ID: 3
    Similarity score: 0.7645
    Research area: Neuroscience
    Result: To better understand animal communication, we must consider all the different sensory modalities involved.
    Reference: 33815100, 9
18. Query ID: 26
    Similarity score: 0.7253
    Research area: Neuroscience
    Result: A channel signal is a linear combination of electrode signals that spatially filters the original multielectrodes recording.
    Reference: 54725834, 36
19. Query ID: 25
    Similarity score: 0.6461
    Research area: Plant Science
    Result: This signaling occurs over timescales that might differ from those traditionally used to measure activity at the neural or behavioral levels and will likely reveal novel and exciting new mechanisms of communication between the body and the brain.
    Reference: 235403109, 153
20. Query ID: 12
    Similarity score: 0.6283
    Research area: Plant Science
    Result: Signalling channels are in charge of regulating these processes, which consist of a reception step (stimuli perception), a transduction step (intracellular and extracellular signal amplification) and, lastly, a response step (enzymatic or non-enzymatic) [43,44].
    Reference: 251201936, 178
21. Query ID: 21
    Similarity score: 0.5607
    Research area: Plant Science
    Result: Auxin signaling is a clear and well-defined pathway; however, its components might be controlled and modified by several other agents such as phytohormones and environmental cues.
    Reference: 254678166, 22
22. Query ID: 4
    Similarity score: 0.7906
    Research area: Psychology
    Result: First, we will summarize the traditional approaches to animal communication, in order of increasing engagement with cognition, and provide the theoretical background to contextualize the model we now propose.
    Reference: 245908229, 33
23. Query ID: 13
    Similarity score: 0.7475

RESEARCH AREA: Psychology
RESULT: Communication involves multiple social cues that are often combined, requiring individuals to process multiple sensory modalities simultaneously.
REFERENCE: 257956451, 28

24. QUERY ID: 17
SIMILARITY SCORE: 0.6985
RESEARCH AREA: Psychology
RESULT: In this context a further organizational aspect deserving consideration concerns the possibility (see Agnati and Fuxe, 2000) that the communication pathway between a signal source and its targets could be under the control of a "modifier," responding to teaching signals that could originate, for instance, from environmental inputs.
REFERENCE: 1724708, 145

25. QUERY ID: 2
SIMILARITY SCORE: 0.7497
RESEARCH AREA: Varia
RESULT: Communication is essential during social interactions including animal conflicts and it is often a complex process involving multiple sensory channels or modalities.
REFERENCE: 33815100, 0

26. QUERY ID: 0
SIMILARITY SCORE: 0.6620
RESEARCH AREA: Varia
RESULT: We will focus on verbal communication, as this is a natural extension of the sender-receiver coupling idea, but in a domain that allows us to establish links between production processes in the speaker's (sender) brain and comprehension processes in the listener's (receiver's) brain.
REFERENCE: 8709249, 74

27. QUERY ID: 7
SIMILARITY SCORE: 0.6458
RESEARCH AREA: Varia
RESULT: The core argument is that, just as a person's utterances reveal only a subset of what they know, animal communication signals express an intrinsically limited subset of that species' conceptual storehouse.
REFERENCE: 208085355, 25

3.2.2 *Identifying the patches of* COMMUNICATION

As even a cursory glance at the examples presented in box 3.2 ascertains, there is indeed a wide spectrum of how the notion of "communication" is used. The literature on patchwork concepts focuses on notions that take multiple meanings within a single field (CORTICAL COLUMN in neuroscience) or in closely related fields (HARDNESS in physics and material science). In these contexts, identifying separate patches is quite straightforward, as the different meanings of the terms are obvious to anyone acquainted with the literature. COMMUNICATION poses special difficulties in this regard. A distinction into possible patches can be

Table 3.1: Statistics of cosine similarities of the results for each query. Confidence intervals have been estimated using Python package `scipy` t-Student distribution model (code included in Supplementary Materials). These intervals can be used to judge the similarity of examples presented in box 3.2 and in the text. All values have been rounded to four decimal places.

Query ID	Mean S_c	Maximum S_c	Median S_c	95% confidence interval
0	0.3704	0.6996	0.3287	(0.3694, 0.3714)
1	0.4139	0.6676	0.3950	(0.4132, 0.4146)
2	0.4152	0.7512	0.3663	(0.4142, 0.4163)
3	0.4418	0.7645	0.4349	(0.4413, 0.4424)
4	0.4644	0.8264	0.4741	(0.4636, 0.4652)
5	0.4665	0.8339	0.4448	(0.4658, 0.4671)
6	0.4292	0.7382	0.3913	(0.4283, 0.4302)
7	0.4340	0.6458	0.4373	(0.4333, 0.4347)
8	0.4690	0.7873	0.4193	(0.4679, 0.4701)
9	0.4116	0.6674	0.3926	(0.4110, 0.4121)
10	0.4339	0.6288	0.4400	(0.4334, 0.4344)
11	0.4302	0.7700	0.3768	(0.4291, 0.4312)
12	0.4577	0.7242	0.4418	(0.4569, 0.4585)
13	0.4015	0.7475	0.3528	(0.4004, 0.4026)
14	0.4249	0.5953	0.4188	(0.4245, 0.4253)
15	0.3917	0.5213	0.3857	(0.3913, 0.3920)
16	0.4157	0.8187	0.3905	(0.4150, 0.4164)
17	0.4116	0.7100	0.3697	(0.4106, 0.4125)
18	0.4544	0.7573	0.4128	(0.4534, 0.4553)
19	0.4378	0.6818	0.4342	(0.4371, 0.4385)
20	0.4792	0.7703	0.4839	(0.4785, 0.4799)
21	0.4247	0.6737	0.4326	(0.4240, 0.4253)
22	0.4471	0.6864	0.4476	(0.4466, 0.4477)
23	0.4341	0.7201	0.4508	(0.4333, 0.4348)
24	0.4408	0.7174	0.4480	(0.4403, 0.4413)
25	0.4744	0.7173	0.4783	(0.4738, 0.4749)
26	0.4606	0.7476	0.4729	(0.4599, 0.4614)
27	0.4565	0.6763	0.4473	(0.4559, 0.4570)
28	0.4333	0.6914	0.4411	(0.4326, 0.4339)

offered by a topic model trained on the results—but to justify the selection of this method, consider two theoretically motivated alternatives.

The first possible strategy comes from the distinciton of "informational" and "influence-based" accounts of communication, as characterized by Scott-Phillips (2008). While Scott-Phillips has an unificational goal of providing a single suitable definition of biological communication, his distinction of the two main approaches is visible in the data: the definitions used as queries can be broadly split into "informational" and "influence." Indeed, they include for instance the standard adaptationist definition proposed by Scott-Phillips (12_BC_DEF1_EX10). However, despite Scott-Phillips' definitive assessment of the latter approach as better suited for explaining what communication is, it is the informational definition (query 4_BC_DEF1_EX2 (mean, median, and max cosine similarity scores, noted here as S_c, all among top 5—see values in Table 3.1), closely followed by queries 20_BC_SIG_EX2, 25_HWB_SIG_EX1 (mean and median S_c among the top 5, with significantly lower values of maximal similarity score) that produce passages with best similarity scores in the corpus (one influence-oriented query closely follows, 5_BC_DEF1_EX3, with the mean and max in top 5, but significantly lower median S_c). One possible explanation could be that the fit of informational definitions stems from the smaller word variation when references to "information" are made (as the word "information" itself does not have any clear synonyms). Indeed, influence-focused approaches allow for greater diversity, as they refer variously to "evolution", "adaptation", "selection" and a plethora of other terms. However, the results produced by querying the corpus with the informational definition do not always mention information directly. Consider the sentence "Communication is essential during social interactions including animal conflicts and it is often a complex process involving multiple sensory channels or modalities." (245908229, 33), which is scored highly by both the informational query 4_BC_DEF1_EX2 ($S_c \approx 0.803$) and influence query 5_BC_DEF1_EX3 ($S_c \approx 0.818$), but does not mention explicitly neither information nor adaptation. Similarly, if we look at the results of querying the corpus with formulations of Scott-Phillips (2008) definitions (0_INFO and 1_ADAPT), we see that while the mean similarity score is higher for the influence definition (judging based on the confidence intervals presented in Table 3.1), the informational one has a higher maximum score, and the two have some overlap with no clear word markers (e.g., the sentence "We will focus on verbal communication, as this is a natural extension of the sender-receiver coupling idea, but in a domain that allows us to establish links between production processes in the speaker's (sender) brain and comprehension processes in the listener's (receiver) brain." (8709249, 74) is picked up by both definitions with scores $S_c \approx 0.662$ for 0_INFO and $S_c \approx 0.64$ for 1_ADAPT). This indicates that while at times the two approaches can be clearly distinguished due to their differing sets of assumptions, this distinction does not clearly define individual patches in the sense of Haueis (2024).

A second possible distinction into patches could be based on the division into the fields of study that are of interest here. This is indeed a quite trivial approach, given that the fields come with their unique spatiotemporal scales, as well as sets of techniques and domains. However, it is highly unlikely that individual disciplines will operate within unified patches. Consider the following sentences attributed to the field of animal communication:

- "Animals may communicate not only about internal states but also about external referents, eliciting also the appropriate reaction from the receiver in the absence of the stimulus triggering the signal in the sender." (205330370, 20)
- "As communication usually occurs in a network of several animals in signalling and receiving range of each other, the emotional state of a caller may influence the behaviour of several individuals, addressees in direct interactions and bystanders alike." (5566073, 24)

and compare them with some others attributed to the same discipline:

- "In the iguana *Liolaemus pacho* for example, chemical signals are thought to have a role in sex recognition (Vicente & Halloy, 2016), while other behaviours like head-bobbing might have a role in attracting the attention of the receiver and communicating the signaller's location (Vicente &~Halloy, 2017)." (165035685, 109)
- "If females prefer leading signals, males may reset their signal rhythm upon hearing their neighbors' signals and are known to restart their own periodic signal to avoid overlapping." (59337737, 16)

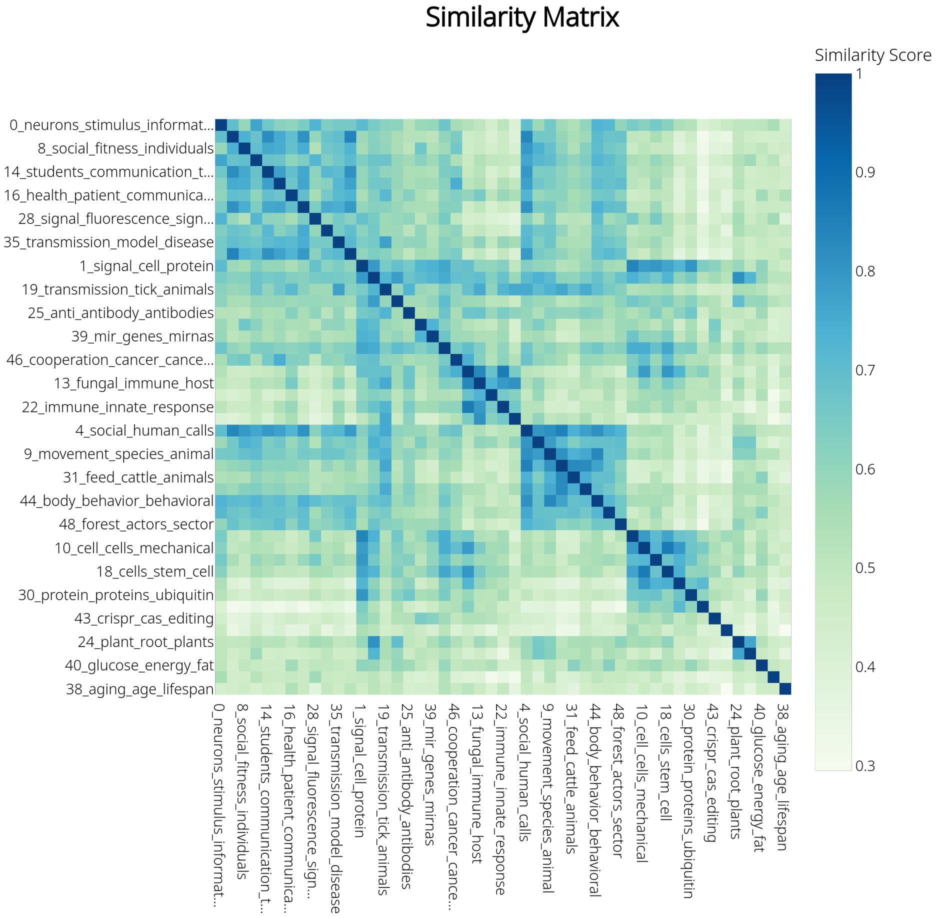
These two pairs of examples differ in the temporal scale which they consider (individual event versus an extended pattern), domain (the second pair considers a specific subset related to sexual signals, while the first one is broader) and the property targeted (while the "technique" remains the same—an ethological study of the behavioural response of the receiver, the property switches from the putative contents of the signal to channel or modality characteristics and role). This suggests that they should be considered independent patches within a single field of study—while not particularly surprising, this shows clearly that fields of study do not approximate patches well enough.

The third possibility—one I ultimately adopt—is to use a topic model of the results as an approximation of patches of COMMUNICATION. This method is data-driven, and focuses on the textual characteristics of the use of the notion. As such, it does not directly depend on the division into fields of study (although, as will become clear, partially recovers it from the data), allowing for identifying patches that recur in different disciplines—e.g., across fields that have some overlap of scale or technique. At the same time this risks omitting the differences between patches which do not have clear local textual markers.

As discussed in subsection 2.5.2, in Step 3. Topic modelling of the analysis the semantic search results have been analysed with the BERTopic algorithm.

I~conducted a manual hyperparameter search to select the appropriate number of topics. The goal was to optimise the coherence-granularity tradeoff: to achieve high coherence scores (using standard measures—C_ν and U_{mass}) while keeping the model within a relatively small number of topics. The search found that the algorithm performs best when the resulting topic model is reduced to 50 topics (49, excluding the “outliers” topic, see the description of BERTopic in subsection 2.2.3). The 10 most popular topics identified in the model are presented in Table 3.2 (this excludes the “outliers” topic which is analysed separately below and ranked as 8th overall in terms of popularity; full list of topics is included in the online supplement).[3]

Figure 3.2: Topic similarity matrix. Topic similarity scores are calculated by the BERTopic algorithm as the cosine similarity between vector embeddings of topics. Interactive version of the figure is available in the online supplement at: https://doi.org/10.17605/osf.io/ax4jm.



---

[3] The online supplementary materials are available at: https://doi.org/10.17605/osf.io/ax4jm.

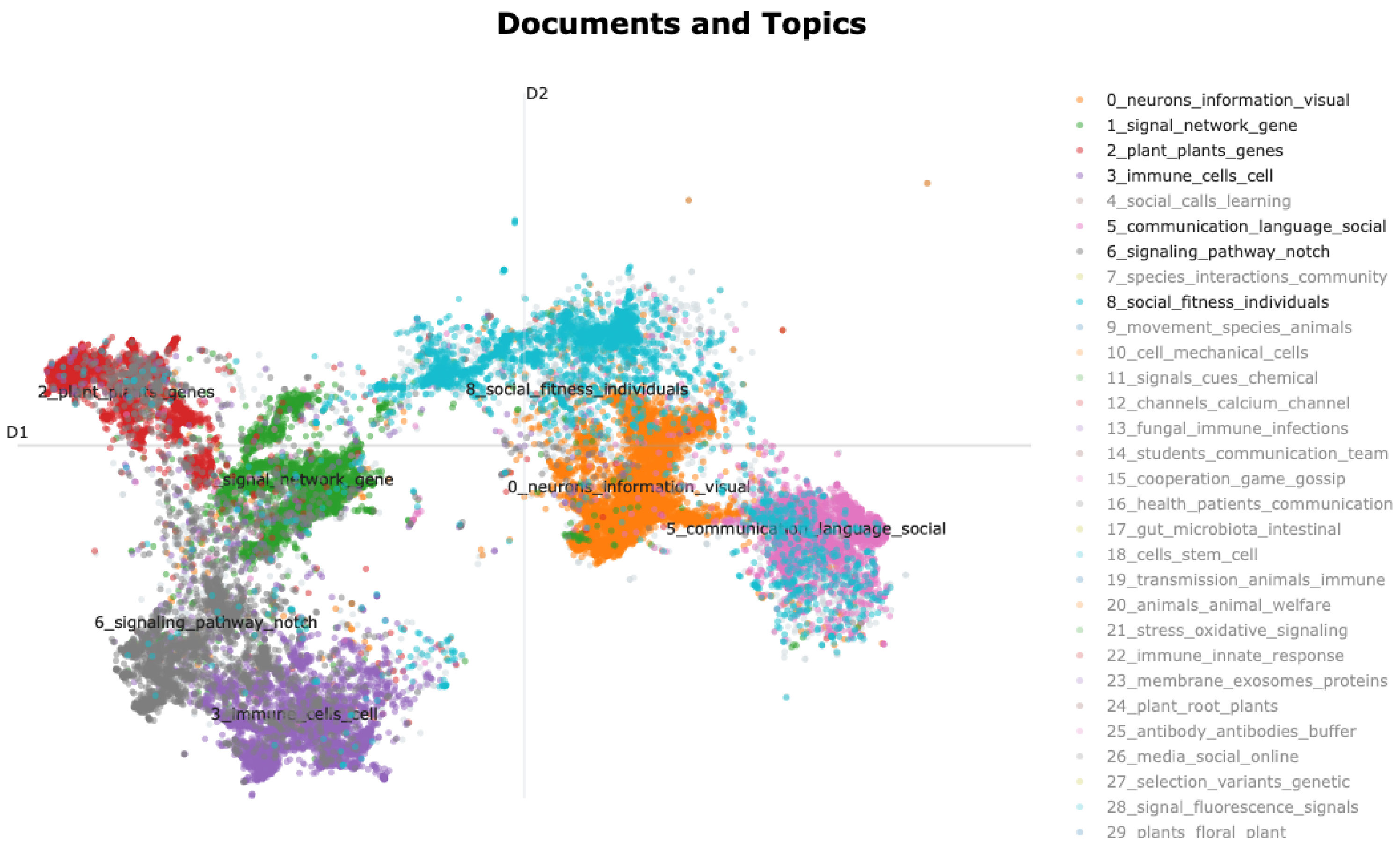


Figure 3.3: Two-dimensional representation of document vector embeddings, calculated by BERTopic using UMAP algorithm. The documents are clustered by topics based on the topic model. Each point represents a 2D-reduced vector embedding of a document (paragraph). Colors correspond to the topics. A subset of most popular topics is selected, for enhanced readability of the figure—the full, interactive version is available in the online supplement at: https://doi.org/10.17605/osf.io/ax4jm.

Table 3.2: 10 most frequent topics with the number of documents classified as belonging to that topic (Count), excluding the “outliers” topic which was ranked 8th overall. I have chosen the topic names through qualitative review of the different representations of the topics produced by BERTopic and a close reading of the representative documents selected by the algorithm (see chapter 2). The table includes all representations used, limited to top 10 key terms for each. Base representation is the default algorithm, the c-TF-IDF (class-based term frequency—inverse document frequency, see subsection 2.2.3 and 2.2.3). “Nouns and modifiers” and “Verbs and modifiers” use Spacy part-of-speech tagger with the `en_core_web_sm` model to extract terms from the c-TF-IDF which match specified syntactic patterns—in this case, nouns, or verbs and their modifiers (if any). “KeyBERT + MMR” calculates the vector distance (semantic similarity) between candidate keywords and topic embeddings, selects best candidates (least distance), and filters them with the Maximal Marginal Relevance algorithm, to minimize redundancy between keywords. Full list of topics is available in the online supplement at: https://doi.org/10.17605/osf.io/ax4jm.

Topic	Count	Name	Base representation	Nouns and modifiers	KeyBERT + MMR	Verbs and modifiers
0	7161	NERVOUS SYSTEM	neurons, stimulus, information, visual, brain	neurons, stimulus, visual, information, brain	stimuli, cortical, neurons, cortex, stimulus	learning, reward, processing, used, using
3	6013	IMMUNE SYSTEM	immune, cells, cell, tumor, immune cells	immune, cells, tumor, cell, cancer	immunotherapy, immune cells, immune response, innate immune, immunity	innate, associated, signaling, related, study
1	5211	QUORUM SENSING	signal, cell, protein, gene, network	signal, cell, protein, network, gene	signaling, gene expression, proteins, bacteria, pathways	signaling, binding, using, used, based

Continued on next page

Table 3.2: (continued)

Topic	Count	Name	Base representation	Nouns and modifiers	KeyBERT + MMR	Verbs and modifiers
2	4634	PLANT PATHOGENS	plant, plants, genes, signaling, defense	plant, plants, signaling, genes, resistance	pathogens, signaling pathways, arabidopsis, biosynthesis, triggered immunity	plant, signaling, pathogen, induced, triggered
6	4241	NOTCH DEVELOPMENT	signaling, pathway, notch, cell, wnt	notch, signaling, pathway, cell, wnt	notch signaling, signaling pathway, notch1, wnt signaling, signaling	notch, signaling, pathways, signalling, activated
5	3785	SOCIAL FACTORS IN LANGUAGE	communication, language, social, children, conversation	communication, language, social, children, conversation	social communication, social interaction, conversations, communicative, autism	talk, use, turn, used, understanding
8	3690	SOCIAL BEHAVIORS	social, fitness, individuals, individual, group	fitness, social, individuals, individual, group	evolutionary, traits, mating, aggression, behavioral	group, given, based, study, used
4	3148	SOCIAL LEARNING	social, human, calls, learning, communication	social, human, calls, dogs, communication	primates, behavioral, chimpanzees, behaviors, species	learning, use, study, used, contexts

Continued on next page

Table 3.2: (continued)

Topic	Count	Name	Base representation	Nouns and modifiers	KeyBERT + MMR	Verbs and modifiers
7	3065	ECOLOGICAL COMMUNITIES	species, interactions, community, communities, microbial	species, interactions, community, microbial, communities	microbial communities, ecological, ecosystem, species, phylogenetic	taxa, used, based, benefit, understanding
20	2182	HUMAN INTERSPECIES INTERACTIONS	animal, animals, welfare, human, dogs	animal, animals, dogs, pets, pet	attachment pets, companion animals, animal welfare, emotional attachment, animal assisted	care, use, stress, assisted, related

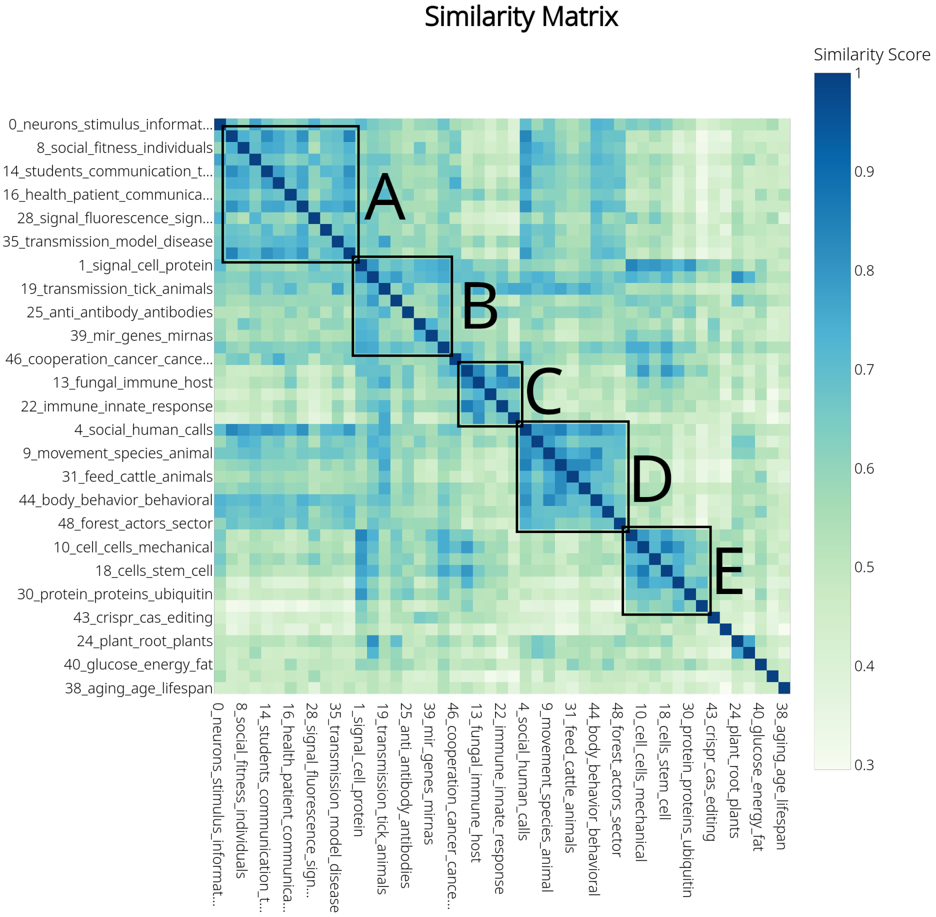
Assessing the topics qualitatively (by investigating the detailed representations of topics and the representative documents selected by BERTopic—partially presented in Table 3.2, full topic model is available in the Supplementary Data), they seem to be distinguished along two dimensions. First, they identify the different subjects that COMMUNICATION is applied to. The three most popular topics refer to different biological systems that extensively rely on electrical or chemical signalling: the nervous (topic 0) and immune systems (topic 3), as well as gene networks which control gene expression (topic 1). Other topics pick out different forms of animal-animal communication (e.g., topics 5, 4, 20). Second, the topics partially highlight the functional role of various communication systems, picking out the involvement of communication systems in cancerogenesis (topic 3), social organization (topics 1, 8, and 4, in different ways) or cognition (topic 0). Together these two dimensions seem to delineate distinct domains of COMMUNICATION. Indeed, the model seems to capture even such nuanced domain distinctions as the ones between topics 2 (PLANT PATHOGENS), 24 (PLANT SYMBIONTS) and 29 (PLANT POLLINATORS). These have quantitatively similar word distributions, as all refer to communication in plants, and specifically between plants and other groups of organisms. However, the topics clearly distinguish between signalling in response to pathogens (topic 2) and signalling as involved in symbiosis (between plants and microorganisms, topic 24, and plants and insects, topic 29). This highlights that the topic model picks out the domain and scale necessary for defining the patches of COMMUNICATION.

The scales for different patches of communication emerge from the qualitative assessment of topic similarities based on the similarity matrix (presented in Figure 3.4, clusters A-E mentioned below are marked in the figure). The topics can be grouped into several "clusters": for instance, cluster A includes topics 5 (SOCIAL FACTORS IN LANGUAGE), 14 (COMMUNICATION IN EDUCATION), 37 (FAMILY COMMUNICATION)—all formulated on the scale of social communication. Similarly, cluster D, which includes topics such as 4 (SOCIAL LEARNING), 20 (HUMAN INTERSPECIES INTERACTIONS), or 9 (ANIMAL MOVEMENT), focuses on the scale of social communication, but extends beyond linguistic interactions. Cluster B, on the other hand, groups topics concerned with the scale of subcellular components involved in signalling—this includes topics 1 (QUORUM SENSING), 21 (PLANT STRESS) or 27 (GENETIC EVOLUTION). A slightly larger scale of cellular processes is grouped by cluster E, which includes topics 6 (NOTCH DEVELOPMENT), 18 (STEM CELLS) or 42 (MITOCHONDRIA). Finally, cluster C captures an intermediate scale of signalling within or between organisms, e.g., in topic 3 (IMMUNE SYSTEM) or topic 34 (PARASITE). This clustering is reflected by the topological structure of the topic space, as represented by UMAP 2D-reduced document embeddings (Figure 3.3), where we see that topics align along a similar structure—from the smallest to largest spatiotemporal scale.[4]

---

[4] Due to the nature of mathematical operations involved in dimensionality reduction with UMAP, and considering its stochastic nature and the impact of hyperparameter choice, the

Figure 3.4: Topic similarity matrix with qualitatively identified groupings of topics marked. Clusters A, D capture the scale of social interactions, B and E—cellular and subcellular signalling, and C groups topics covering signalling within or between organisms. For fuller description see the text.



In fact, the topic model, at least in some cases, provides us directly also with the other two components necessary to define the patches, the techniques, and the properties they target—included either in the representation of the topic produced by BERTopic, or in representative documents. These sometimes occur separately, like in the case of properties without mention of techniques: vocalizations (topic 4), specific chemical substances or biochemical compounds, including calcium (topic 12), pheromones (topic 11), proteins such as Wnt or Notch [5] (topic 6). However, they also come packaged together: as in the case of questionnaires (technique) of relationship satisfaction (property; topic 37), models (technique, albeit quite a broad one) of the dynamics of disease spread (property; topic 35), and genome-wide associations (technique) of phenotypic traits (property; topic 27).

Box 3.3: Patches of COMMUNICATION corresponding to the 10 most popular topics identified in the dataset.

In some cases, techniques and properties were supplemented based on qualitative analysis of articles assigned to this topic and broader domain knowledge, as these elements were not always readily available from the topic representation and from representative documents, as picked out by BERTopic.

1. TOPIC ID: 0
   TOPIC NAME: NERVOUS SYSTEM
   SCALE: Multiple interacting cells
   DOMAIN: Neurons, neural populations (cortex, brain)
   TECHNIQUE: Electrophysiological and imaging methods: single cell recordings, electrical stimulation, two-photon calcium imaging etc.
   PROPERTY: Activation pattern

2. TOPIC ID: 3
   TOPIC NAME: IMMUNE SYSTEM
   SCALE: Multiple interacting cells
   DOMAIN: Immune system: immune and cancer cells, their components, incl. antibodies, systemic and local properties (genetic factors, biochemical compounds and their networked interactions), on the timescale of tumor development
   TECHNIQUE: In vitro models (e.g., 3D microfluidic devices), flow cytometry, proximity-

---

structure the visualization highlights needs to be interpreted cautiously—hence, here I rely primarily on the parallels between structures highlighted by UMAP and topic similarities.

[5] Notch signalling pathway, to which I will refer in various places throughout this thesis, is a conserved cell signalling mechanism, playing a major role in development (regulating cell proliferation, differentiation, and death) and its misfunction is implicated in a number of diseases, including many cancers (see the reviews in Kopan 2012; Zhou et al. 2022). It's been discovered over a century ago and is an important subject of contemporary biological research. The mechanism consists of a Notch protein, with an intra- and extracellular domain. Roughly speaking, in normal circumstances, upon direct cell-to-cell contact with a cell expressing an appropriate transmembrane ligand, the extracellular domain of Notch binds to the ligand, leading to the cleavage of the intracellular domain, which eventually enters cell nucleus and modifies gene expression (furtherproteins are involved in this process, regulating the signalling activity).

based intercellular labeling (e.g., LIPSTIC), RNA sequencing
PROPERTY: Tumor proliferation, patient survival

3. TOPIC ID: 1
   TOPIC NAME: QUORUM SENSING
   SCALE: Multiple interacting cells
   DOMAIN: Bacterial biofilms, colonies, cell signalling mechanisms (genes, proteins)
   TECHNIQUE: Microscopy, calcium fluorescent imaging, genetic probing, mathematical modelling (differential equations)
   PROPERTY: Biofilm growth, cell behaviours, signalling molecule densities

4. TOPIC ID: 2
   TOPIC NAME: PLANT PATHOGENS
   SCALE: Individual cells
   DOMAIN: Receptors, genetic and protein mechanisms within plant cells
   TECHNIQUE: (unclear)
   PROPERTY: Cell behaviours (e.g., programmed cell death, biochemical changes: protein expressions, hormone production etc.)

5. TOPIC ID: 6
   TOPIC NAME: NOTCH DEVELOPMENT
   SCALE: Individual cells
   DOMAIN: Morphogenesis, cell development and specialization and their subcellular mechanisms
   TECHNIQUE: Immunofluorescence, immunoblotting, genetic methods (reporter lines, knockouts), flow cytometry
   PROPERTY: Gene expression, cell viability and specialization

6. TOPIC ID: 5
   TOPIC NAME: SOCIAL FACTORS IN LANGUAGE
   SCALE: Multicellular organisms
   DOMAIN: Verbal linguistic communication: capacities and use
   TECHNIQUE: Neuropsychological methods (e.g., study of aphasia patients), questionnaire-based methods, psychological experiments
   PROPERTY: Vocabulary, language processing efficiency

7. TOPIC ID: 8
   TOPIC NAME: SOCIAL BEHAVIOURS
   SCALE: Multicellular organisms, multiple interacting organisms
   DOMAIN: Group behaviours, social behaviours of individuals
   TECHNIQUE: Ethological studies, mathematical models
   PROPERTY: Behavioural patterns, fitness

8. TOPIC ID: 4
   TOPIC NAME: SOCIAL LEARNING
   SCALE: Multiple interacting organisms
   DOMAIN: Social transmission of behaviour
   TECHNIQUE: Ethological studies
   PROPERTY: Behavioral repertoires of individuals and populations

9. TOPIC ID: 7
   TOPIC NAME: ECOLOGICAL COMMUNITIES
   SCALE: Multiple interacting groups
   DOMAIN: Ecological species interactions

Technique: Microscopy, mathematical models
Property: Diversity, fitness

10. Topic ID: 20
    Topic Name: HUMAN INTERSPECIES INTERACTIONS
    Scale: Multiple interacting organisms
    Domain: Human attitudes and behaviours towards other groups of animals
    Technique: Questionnaire-based methods, case studies
    Property: Human opinions, human welfare

To substantiate this approach, let's look at the patches corresponding to the 10 most popular topics, presented in box 3.3. While the topic model on its own does not provide all the necessary elements for the definition of all the patches, it clearly picks out different classes of phenomena that are indeed often discussed in terms of communication. These could be further divided into conglomerates of different, but largely overlapping patches, as in the case of the NERVOUS SYSTEM topic, which includes, among others:

- Communication$^{\text{single neuron}}$⟨ two-photon calcium imaging, nervous system(calcium concentration) ⟩
- Communication$^{\text{single neuron}}$⟨ single-cell recordings, nervous system(electric potentials of cell membrane) ⟩
- Communication$^{\text{voxel}}$⟨ functional magnetic resonance imaging, nervous system(temporal BOLD correlation) ⟩

These patches are connected not only by the domain but also, most importantly, by the theoretical assumption of causal dependencies or constitution between the chemical and electrical properties of cells and the activity of individual neurons and whole neural populations. In result, it is taken for granted that functional magnetic resonance imaging (fMRI) and single cell recordings ultimately measure phenomena of the same class. Similarly, it is believed that the results of these different methods can be connected with the help of, e.g., computational modelling (but then again—modelling necessarily involves decisions regarding scale). Such connections are much less obvious when we consider the patches from a different domain.

Consider the patch identified by the topic IMMUNE SYSTEM, which similarly consists of multiple overlapping patches:

- Communication$^{\text{single cell}}$⟨ proximity-based intercellular labeling, immune system(peptide tags labeling interactions) ⟩
- Communication$^{\text{subcellular}}$⟨ RNA sequencing, immune system(expressed proteins) ⟩

These two subpatches are again quite strictly (causally or constitutively) connected: cell-to-cell interactions are governed by the mechanics of subcellular components, and in turn constrain those lower-level processes (e.g., by modifying the available substrates). Despite the difference in scale, the COMMUNICATION involved is clearly of the same kind.

Finally, to shift the scale significantly, consider the patch identified by the topic SOCIAL FACTORS IN LANGUAGE:

- Communication$^{\text{human}}$⟨ questionnaire-based survey, verbal language(vocabulary size) ⟩
- Communication$^{\text{human}}$⟨ psychological experiment, verbal language (language processing efficiency - response time) ⟩

This topic quite clearly identifies a single patch, which covers the varied cases and phenomena associated with verbal linguistic communication: starting with its acquisition from a developmental perspective (and the impact of socioeconomic factors on language learning), via its online execution (and the impact of social factors, such as gaze direction of the interlocutor), to its breakdown (e.g., in aphasia). Studying these phenomena will involve different techniques and lead researchers to focus on varying properties, but does not lead to any ambiguity in the notion of COMMUNICATION.

While some topics identify different patches, there is also some overlap of patches between the topics. Consider the Communication$^{\text{subcellular}}$ patch picked out by the IMMUNE SYSTEM topic, and the patch identified by the topic PLANT PATHOGENS:

- Communication$^{\text{subcellular}}$⟨ RNA sequencing, plant cells (expressed proteins)⟩

These two are closely connected, despite the difference of domain. While plants do not have a separate, adaptive immune system based on mobile defender cells, like the one found in many animal groups, individual cells can exhibit an innate immune response to pathogens, and the molecular mechanisms of that response are often shared over the wide gap in phylogeny. An example is the expression of a single type of enzyme, the mitogen activated protein kinase, in response to infection. These kinases are involved in critical signalling pathways that are studied using analogous techniques across those different domains, including RNA sequencing (Elion 2006). Hence, the COMMUNICATION involved in immune responses, whether in plants or animals, seems to constitute the same patch.

3.2.3 *Connecting the patches*

The idea of “patchwork concepts”—specifically, the presumed prevalence of local over global relations between the different uses—assumes that the groupings of phenomena captured by individual patches will be dissimilar and difficult to map between the domains. In analysing the patchwork structure of the term HARDNESS, Mark Wilson observes (M. Wilson 2006, p. 338): “As an inferential tool, most of us employ the term ‘hardness’ informally as a generic, single-scale ‘folk physics' parameter that we consult in selecting a material for, e.g., manufacturing purposes. Left at this general level, the term requires further specialization before it can carry much data usefully." What he means is that there is an inherent polysemy in the general English notion of "hardness". This polysemy allows us quite easily to distinguish between the hardness of different materials. The hardness of materials in a smartphone touch screen can be considered by its resistance to scratches (think only of those soft, plastic screens so common among smartphones a few years back, and the currently more prevalent glass screens). The hardness of a shoe sole, can be investigated by bending the shoe on our foot or by jumping to observe the dynamics of the indentation of the material (think of the hardness we want from a heel of high, formal heels and the dynamics we expect from cushioned running shoes). We consider these to be different aspects of an analogous property, as exhibited in very different materials, but the actual physical dependencies of these different notions only become apparent through the complex dependencies of various operationalizations of HARDNESS and the domains they apply to.

As a second step of the argument by contradiction developed in this chapter, I would like to look carefully at the emerging structure of those patchy operationalizations of COMMUNICATION. The goal is to establish whether this structure is fully accounted for by local relationships between individual patches, or whether an underlying core meaning shared by all patches accounts for those connections better.

To begin, note that there are clear local relations emerging from the previous characterization of patches: the PLANT PATHOGENS and IMMUNE SYSTEM patches differ only relative to how we conceive of immune systems, ultimately constituting a single patch. The IMMUNE SYSTEM patch has a local relationship with the QUORUM SENSING patch, as both rely partially on identifying various protein kinase signalling pathways for mapping the actual mechanisms of cell-to-cell communications. The QUORUM SENSING patch then connects to the NERVOUS SYSTEM patch partially by their joint reliance on the same techniques (calcium fluorescent imaging) and partially due to the shared properties—the activation patterns of neurons being a particular, specialized form of the biochemical changes in bacteria, enabled by the activity of ion channels. Across scales there is a link between the QUORUM SENSING and SOCIAL BEHAVIOURS patches: while the group behaviour of social insects will be different in critical respects from that of colonial unicellulars, to mention only the functional specialization (although relevant intermediary cases do exist, such as the *Volvox* algae, which is studied as a form of proto-multicellularity, e.g., Herron 2016), the two patches are connected by their focus on how the individual interactions impact the group organization and, vice versa, how the social organization impacts the individuals—even if the signalling which mediates these patterns of behaviour will exhibit very different complexity. This then leads to a strong similarity between these two patches and the one identified by SOCIAL FACTORS IN LANGUAGE, despite no clear connections between their defining elements.

Interestingly, these relations—which emerge from the analysis of the documents assigned to these topics—are not immediately supported by the topic similarity scores (the similarities are encoded in the colour intensity of the appropriate square in Figure 3.2, exact numerical values are available in the interactive version in the online supplement). For example, the topics IMMUNE SYSTEM and PLANT PATHOGENS have a similarity score of approx. 0.69, higher than average similarity for all topics (approx. 0.56, with a 95% confidence interval of approx. $(0.55 - 0.57)$, calculated from a t-Student distribution), but lower than the highest similarity score of PLANT PATHOGENS and QUORUM SENSING topics (approx. 0.79) or for IMMUNE SYSTEMS and INNATE IMMUNITY ANIMALS (topic 22, approx. 0.86) or FUNGAL INFECTIONS (topic 13, approx. 0.83), which are all concerned with animal immune systems. IMMUNE SYSTEMS, FUNGAL INFECTIONS and INNATE IMMUNITY ANIMALS refer to patches which are closely connected through the shared domain. The connection of PLANT PATHOGENS and QUORUM SENSING is more tenuous, and the high topic similarity score is likely a result of the shared vocabulary referring to bacteria. These are the subject of study in both cases, although from quite different perspectives (note that this overlap in vocabulary is not present for most of the topics which do not have a clear local relation). Nevertheless, this similarity indicates close connections between the techniques and properties used in the two contexts.

These local ties construct a multidimensional structure, which highlights an important feature of patch relationships more generally: the local connections are not transitive. Consider the relation between the IMMUNE SYSTEM and NERVOUS SYSTEM patches, which share QUORUM SENSING as a common neighbour (in the sketched patch structure suggested above). While both are concerned with the same scale of multiple interacting cells, the differences in the characteristics of the domain result in significantly different techniques and properties targeted. One of them is due to the dynamic character of the immune system, compared to a relatively stable structure of the nervous system: techniques such as proximity-based intercellular labelling do not make much sense for studying the connections in the nervous system, which instead relies on electron microscopy of dissected and stained brain tissue for studying physical interactions between cells (or on fMRI and calcium imaging techniques for functional connectivity). Effectively, such relatively marginal differences accumulate, leaving the patches to be clearly delineated, despite the shared neighbour.

3.3 TRIANGULATING COMMUNICATION

With the patches identified and an emerging view of their relations, we can turn to the question of whether the definition-schema can be considered the general reasoning strategy. In this section, I will highlight that the multidimensional structure of the relations between "patches" of communication escapes the characterization in terms of a patchwork, and introduces a core meaning of COMMUNICATION.

Haueis (2024) notes that the generalised reasoning strategy is what justifies the extension of the concept in question to novel cases and constitutes an "underspecified core meaning" (p. 746) of the patchwork. As such it can be uncovered both synchronically, by matching the similarities in uses of the concept (e.g., the concept of HARDNESS, where the reasoning strategy can be explicated as instructing "researchers to, first, perform a mechanical intervention on a material and, second, find a quantity that describes how the material resists that intervention", Haueis 2024, p. 746), or diachronically, by tracing the shifts that the concept has made (e.g., the concept of CORTICAL COLUMN, which has been initially formulated at the macroscale (Hubel and Wiesel 1977) and further extended to meso- and microscale as novel techniques permitting the investigation of smaller-scale structure have been developed (Mountcastle 1997)).

While a detailed historical investigation of the traverses that COMMUNICATION has undergone goes beyond the scope of this dissertation, a few general points should be noted, to justify the overall direction of the subsequent argument. In his history of early neuroscience on the Indo-European continent, Stanley Finger (2001) includes quotes from Galen that indicate that already the early recognition of the brain engagement in cognition involved notions of "receiving" (e.g., "in substance the encephalon is very like the nerves [...] and this was proper for a part that was to receive all sensations" Galen 1968, p. 398; quoted in Finger 2001, p. 16), as well as the idea of "flows" and "travelling spirits" (Finger 2001, p. 16-17). This notion of flows and a broader hydraulic perspective on bodily function, including the brain, has persisted over several centuries, most importantly guiding Descartes' ideas of the brain (see Descartes [1667] 2018). It was finally overhauled with the discovery of nerves' electric conductance that begun in the second half of the 1700s (Cobb 2020, 2021) and through the advancements of Galvani, culminated in the work of Hermann von Helmholtz that laid the foundations for much of contemporary neuroscience. It is at this point the notions of flows, transfers and connections, present previously, translate quite naturally—at least for the 1800s researchers—into the metaphor of the nervous system as a telegraphic device. Matthew Cobb ascribes the origin of this notion to Alfred Smee (Cobb 2020, chapter 3) and notes a slightly later quote from the 1853 biological dictionary that explicitly uses the term "information": "the wires – nerves – convey the information from all parts of the body" (Thomson and H. H. Smith 1853; quoted in Cobb 2021). Ultimately, the contemporary dominance of this conceptualization results directly from the impact of Shanonn's mathematization of information and the resulting formal description of communication systems (including its uptake in game theory, see Lewis 1969, explored in detail in chapter 4). The overall direction is pretty clear: macroscale, highly technological areas of contemporary best science are applied to the understanding of mental function, and as a result, translated to the meso- and microscale. All the listed historical perspectives include some notion of transfer or connection, and while in some cases these are more mechanical than communicative (e.g., the view of vibrations travelling through nerves proposed by Jan Swammerdam, [1676] 1758; as discussed in Cobb 2021), they eventually lead to the informational approach.

The notions of "transfer" or "connection" are initially applied to describe the activities of the organ which is considered the biological seat of the mental—whether it was the heart, or later the brain and the nervous system. At times, these notions get further translated into other biological areas. This could be considered a part of a broader methodological anthropocentrism of sorts: consider, for example the similar type of reasoning at play when social structure of animals or bacteria is the object of study—"animals serving as surrogates for 'man' ", as Donna Haraway puts it in her discussion of primatology (1992, p. 2). This anthropocentrism could be cast as a metaphorical or analogical reasoning, where the unknown is explained through the known by picking out similarities.

Metaphors serve a variety of different functions in scientific practice (see Boyd 1993). Analogous reasoning, especially as applied to structural properties, emerges historically as the more austere and firm version of metaphorical reasoning, that fits within the strict constraints of the logic of contemporary science (see Gentner and Jeziorski 1993; for central account of analogy in science see Hesse 1966; Hesse 2017). The historical trajectory of communicative notions seems to follow a shift from a looser, metaphorical relationship to a narrowed analogy, that ultimately led to the construction of the more abstract mathematical notion of communication. This notion, in turn, underscores a shared organization of the processes it is applied to.

More generally, communicational notions were originally introduced (quite obviously) for the domain of human language and social interactions. From there, they have been applied to descriptions of biological function, primarily of the nervous system. What Smee's telegraphic metaphor shows, is that this expansion of scope has relied both on the structural similarity of the nerves to telegraph wires, but also on the functional parallels, where the excitations travelling along the nerves convey—and integrate—information (note the imperial thinking embedded in the Thomson quote above, with one-directional transfer of information from the periphery to the centre).

In relation to the question at hand, this historical trajectory motivates us to search for the general reasoning strategy starting from the scale of intentional (or ostensive) communication and from (human) language. Notably, the preference towards this scale is quite clear in the preceding discussion of existing definitions of BIOLOGICAL COMMUNICATION—how they are restricted to the "animal communication" (which includes the patches SOCIAL BEHAVIOURS and SOCIAL

LEARNING, among others). Indeed, intentional communication has been the subject of a great deal of scholarship. But while there are different competing approaches (see the review in Scott-Phillips 2025)[6], most work on intentional (or ostensive) communication highlights the three core conditions I have already identified above:

1. the involvement of two distinguishable entities, the sender/communicator and the receiver/audience;
2. the transfer of "something", whether it's a message encoded into signal or a piece of evidence for informative intention;
3. the existence of a structure that allows this transfer: for example, in the form of a shared code or the presence of appropriate intentions on both communicating sides.

Depending on the particular approach we adopt (some version of the "code model" or "ostensive model", as distinguished by Scott-Phillips (2025)), the more generic components of my proposed definition-schema are here completed with particular implementations which have differing connotations (e.g., "signal" vs. "informative intention", see subsection 5.4.1). Nevertheless, they share the same organization.

3.3.1 *General reasoning strategy for* COMMUNICATION

As the emerging picture shows, the possible general reasoning strategy for COMMUNICATION relies on the imputation of a particular organization to the processes analysed in terms of COMMUNICATION. This organization, as I have already discussed (subsection 3.2.1), is in fact shared across most of the existing theoretical accounts:

(1) transfer of a *certain mark* K,
(2) involvement of *distinguishable* entities, X and Y, and
(3) the presence of a *structure* S enabling such transfer.

Further, communication is usually considered as embedded within a broader system, where it serves a particular function—an obvious statement, but one that highlights that instances of communication will be identified as involved in other processes. This can be considered a fourth component of the putative general reasoning strategy, one that has been overlooked by the theoretical accounts discussed before:

(4) The functional role of a communicative process.

---

[6] Scott-Phillips presents this work falsely as dichotomous, focusing on different models of semantics. I dive into this topic in greater detail in chapter 5, and discuss Scott-Phillips' model in section 5.4

Consider the passages in box 3.4, picked out as representative for a given topic by the model introduced above, where the majority highlights this functional nature of communication.

Two important issues need to be noted. First, while only few of the sentences mention "communication" directly, the majority does directly describe a variety of communicative processes, relying instead on notions of "signal" and "signalling." These seem to be more directly tied to the processes occurring at the molecular or single-cell scale, however the sample does not indicate a clear distinction, with one of the passages stating that *signals mediate cellular communication* (see example for topic 6 in box 3.4). I return to this distinction more systematically in section 3.4.

Second, interestingly, some passages in box 3.4 (especially examples for topics 7, 8, and 20) make little use of explicit communicative notions. Instead, examples for topics 7 and 8 describe various forms of interactions ("within-group responsiveness" or "interspecific interactions of locally linked communities") that—given the employed methodology—can be assumed to share at least some properties with communication.[7] Indeed, the example for topic 20, on the other hand, likely groups communicative and non-communicative interactions together in virtue of their effects, e.g., their impact on care, attachment, wellbeing. Nevertheless, it achieved high scores in relation to multiple queries.[8] From this perspective, we can interpret this sample sentence as highlighting the long-term benefits of a sustained, social communication system.

Box 3.4: Sample sentences for topics, selected from the most popular topics.

Sentences were selected through manual, qualitative exploration of the paragraphs marked as representative documents by the BERTopic algorithm. The goal was to highlight how the different topics use the term "communication", focusing on the most explicit uses of the term.

1. Topic ID: 0
   Topic Name: IMMUNE SYSTEM
   Document: The sensitized TLRs trigger a cascade of signaling that activates IRFs and NF-κB leading to IFN (type I) and proinflammatory cytokines production.
   Reference: 237593934, 55
2. Topic ID: 0
   Topic Name: IMMUNE SYSTEM

---

[7] In fact, the sample sentence for topic 8 presented in the table was not found during semantic search, but belongs to the context of another sentence, which scored ≈ 0.45 on query 15_HWB_EX1, significantly above the 95% confidence interval (approx. $(0.391, 0.392)$, as indicated above in Table 3.1). The sample sentence for topic 7, also comes from a context of one of the query results, which had the highest score of ≈ 0.486 to query 8_BC_DEF1_EX6 (confidence interval of approx. $(0.468, 0.470)$).

[8] The best similarity score was achieved in response to the query 3_BC_DEF1_EX1—≈ 0.576, with the 95% confidence interval $(0.4413, 0.4424)$.

Document: MAPK transforms extracellular signals into intracellular signals and exerts biological effects by regulating the expression and function of relevant genes and proteins.
Reference: 232265138, 202

3. Topic ID: 1
   Topic Name: QUORUM SENSING
   Document: Signal molecules increase in secretion as the density of bacteria increases. When a signal molecule reaches a certain threshold, the signal molecule binds to the corresponding signal molecule receptor and activates the receptor. The activated receptor then activates the relevant transcriptional regulators to synthesize extracellular polysaccharides, toxic factors, alginates, etc., thereby causing bacteria to form biofilms.
   Reference: 213192762, 149

4. Topic ID: 1
   Topic Name: QUORUM SENSING
   Document: This process [i.e., quorum sensing] involves the production and detection of signaling molecules (called autoinducers) allowing bacterial communities to express genes collectively.
   Reference: 3774833, 214

5. Topic ID: 2
   Topic Name: PLANT PATHOGENS
   Document: Once PTI [PAMP-triggered immunity] is activated, a series of signaling takes place to counter pathogen attacks, i.e., stomatal closure to prevent pathogen invasion, cell wall thickening and lignification or callose deposition, ion fluxes and oxidative burst; release of reactive oxygen species (ROS), synthesis and release of defense-related hormones; ethylene and salicylic acid.
   Reference: 253372539, 23

6. Topic ID: 3
   Topic Name: NERVOUS SYSTEM
   Document: These studies suggest that signals carrying information about behavioral outcome propagate back to early sensory cortex, and that this "top-down" influence (e.g., expectation or prediction) sets cortical state in preparation for the ongoing behavioral demand.
   Reference: 52938472, 93

7. Topic ID: 4
   Topic Name: SOCIAL LEARNING
   Document: However, most primate studies focus mainly on foraging techniques, and communicative signalling has been overlooked. Some cases of culturally transmitted communication behaviour concern vocal traditions (e.g. nestbuilding calls by free-ranging orangutans, Pongo pygmaeus; Wich et al., 2012) and visual signalling traditions associated with vocalisations (e.g. howler monkeys, Alouatta pigra, placing their hand in front of their mouth while vocalising; Briseño-Jaramillo, Estrada & Lemasson, 2015). Other examples include gestural communication traditions [e.g. HANDCLASP GROOMING by chimpanzees (Nakamura, 2002); HAND EXTENSION by mandrill, Mandrillus sphinx (Laidre, 2008)].
   Reference: 209416584, 17

8. Topic ID: 5
   Topic Name: SOCIAL FACTORS IN LANGUAGE
   Document: Social interaction requires the ability to distinguish between communicative and non-communicative gestures, which involves identifying communicative

intent. This was afforded in the social condition given that non-communicative and communicative eye movements were intervened by eye-contact; an ostensive signal for communicative intent.
Reference: 265364892, 163

9. Topic ID: 6
   Topic Name: NOTCH DEVELOPMENT
   Document: Notch signaling is activated either as paracrine signal to mediate communication between two different cell types or as molecular event involving stem cells differentiation.
   Reference: 231149762, 73
10. Topic ID: 7
   Topic Name: ECOLOGICAL COMMUNITIES
   Document: Here we employ a spatially explicit metacommunity model to investigate the effect of local dispersal on the structure and function of communities in which species are linked through an interaction web comprising mutualism, competition and exploitation. Our results show that function, diversity and interspecific interactions of locally linked communities undergo a phase transition with changes in the rate of species dispersal.
   Reference: 14105589, 1
11. Topic ID: 8
   Topic Name: SOCIAL BEHAVIOURS
   Document: Within-group responsiveness between neighbours allows for cohesion and consensus to form in movement decisions [2], predator avoidance [3] and resource exploitation [4], and thus provides important benefits for individual group members [5,6].
   Reference: 1308391, 16
12. Topic ID: 20
   Topic Name: HUMAN INTERSPECIES INTERACTIONS
   Document: In recent years, the bonds formed between humans and their pets have more frequently been regarded as attachment relationships that often benefit both the animal and human in terms of health and welfare outcomes.
   Reference: 195890915, 59

With that in mind, we can look for the suggested elements of organization in the examples from box 3.4:

- Component (1), transfer of a mark, is mentioned directly through the reference to "signals", "signalling molecules", "gestures", "calls" in the majority of included examples.
- Component (2), the distinction of sender/communicator and receiver/audience, is again quite universal, though specialized to the domain: "early sensory cortex", "molecule receptors", "cells", "neighbours."
- Component (3) is perhaps the most elusive: there are some mentions of "tradition" to highlight the historical embeddedness of the signals, or of "intent", which indicates a particular expectation of the receiver. Other topics mention that the signals are part of a series of molecular events (which touches more upon function, discussed below), or specific molecular components that allow for the picking up of signals, like "MAPK." Most of

the sentences, however, refer to structure indirectly: through verbs such as “propagate”, “trigger”, “link”, or through notions of “threshold” and the identification of a particular architecture, e.g., “top-down.”
- Component (4), the reference to the function that communication serves is, on the other hand, almost universal: communication leads to the “preparation for the ongoing behavioral demand”, “proinflammatory cytokines production”, “expression of relevant genes and proteins”, including collectively, “forming biofilms”, “prevent pathogen invasion”, “stem cell differentiation”, “social interaction”, “consensus in movement decisions”, “phase transition in function, diversity and interspecific interactions”, or “benefits in terms of health and welfare.” Importantly, most of the highlighted examples underscore the fact that communicative instances rarely occur in isolation, but rather are grouped into pathways, series, or cascades of signalling events. This is an important notion and highlights limitations of some intuitions regarding communication that I turn to in the next section.

Both organization and function are also highlighted by the characteristic collocations of “communication” in the results of semantic search, as identified by the Word Sketch algorithm (Step 4. Word Sketch, see subsection 2.5.2):

- Component (1) is referred to mainly through modifiers, such as “verbal”, “non-verbal”, “gestural”, “chemical”, “multimodal” or “vocal” communication.
- Component (2) is mentioned both by modifiers (“intercellular”, “interpersonal”, “cell” communication) or by nouns and verbs that “communication” modifies (communication “partner” or “network”).
- Component (3) is most directly included in the nouns and verbs modified by “communication”: communication “strategy”, “style”, “method” or “capacity”, which indicate the constraints imposed, communication “system”, “channel”, “technology” or “tool” which highlight the materiality of the structure involved. But it is also highlighted by the verbs that take “communication” as an object: they indicate that communication needs to be “mediated”, “facilitated”, “enabled” or “enhanced”; as well as verbs that take it as a subject: communication “requires”, “relies”, “depends” or “needs” a number of preconditions to be met. This spotlights the constraints on the structure of communicative processes—an issue I will discuss at greater length in the next chapter (chapter 4).
- Component (4), the function, is indicated by the verbs that have “communication” as a subject: communication can “help”, “play”, “provide”, “allow”, “serve” or “influence” other processes. As such, adjective predicates of “communication” provide an evaluative framework within which communication can be regarded as “important”, “essential” or “crucial” but also “effective” or “successful.” These are quite generic and do not indicate any specific functions, but rather the broader involvement of communication in a broad variety of processes.

The proposed general reasoning strategy does indeed apply quite well to the examples of the use of communicative notions—including contexts where communication seems to be only implied. This is however not *just* a generalised reasoning strategy: it defines a core meaning of the concept. The strategy itself, possibly due to its focus on the causal organization of biological phenomena, does not require additional specification in a technique-involving way. Whether we mean two adult humans talking to one another or signalling between two cells of a regenerating tissue in a frog limb, we seem to be referring to the same type of process. The apparent polysemy can instead be attributed to the four components which themselves make use of polysemous, patchwork concepts that require specification in a technique-involving way: the notions of transferred marks, or of individual participants of the interaction, do not have a clear meaning and seem to be indexed to the domain under consideration (think of the neurons in the brains of those talking adults). And yet, the broader concept of COMMUNICATION abstracts those ambiguities away, suggesting that there is in fact a shared, scale-free, core meaning of COMMUNICATION captured by the definition-schema advanced above with the fourth, functional component.

The analysis undertaken here in fact arrives at a similar position to the one formulated by Arnon Levy (2011) in his discussion of the use of the notion of "information" in biology (covered in more detail in chapter 4). In that context, Levy (2011) advocates for a form of fictionalism with regard to biological information—primarily understood as semantic (see also the discussion in chapter 5). In his view, information acts as a "liminal metaphor", where its metaphoric character is backgrounded to highlight the epistemic benefits it brings. But it is metaphoric nonetheless: Levy notes the epistemic hedging of mentions (e.g., the use of 'scare quotes') to semantic properties in the literature and supports this argument by stating that this is "the most natural way to take" information talk (Levy 2011, p. 647) given the widespread use of information across many domains of biology. While perhaps not the strongest argument, this viewpoint is quite understandable, especially when we consider the indirect way semantic ascriptions are made in the literature (see the data analysed in chapter 5). Crucially, as Levy's analysis highlights (2011, pp. 649-650), the metaphor of biological (semantic) information can be beneficial because the information talk picks out a particular causal pattern characterized by:

1) directionality—the separation of sender and receiver (spatially or temporally) often with a boundary, and the associated direction of influence;
2) connecting variation—information is invoked to highlight systematic co-variation of two components, especially when they're connected through a complex mechanism with many intermediate links;
3) activity of the sender and receiver—though this is metaphorical according to Levy, informational language allows for specifying a particular component of the process as the active sender or receiver, and another—as a passive signal.

This pattern corresponds to the proposed components of communication in an interesting way. It presumes the distinction of a sender and receiver, and introduces some additional restrictions into what marks and structures can support communication. Crucially, Levy derives these conclusions by tracing how *information* is used in biology, in particular in the context of debates described in the Introduction (chapter 1). The match between the causal structure Levy proposes as the target of informational metaphors, and the communication structure I have identified, suggests that indeed the biological talk of information reviewed previously in the section 1.3 is intended to capture an overlapping set of biological phenomena. I will go back to the context of Shannon's formalism in the next chapter (chapter 4) and the semantic aspect in chapter 5.

More immediately, as my analysis in this section indicates, the three properties noted by Levy can also be applied to non-semantic references to communication. As already mentioned, the shared causal structure, captured by the definition schema and further specified with empirical data, can be adopted in any of the contexts that COMMUNICATION is used, and it seems that this is what motivates the extension of the notion to various domains. But it is more than a generalised reasoning strategy, as the description applies directly to those contexts, without the need for specifying the components in a technique-involving way—a point that will be relevant also in subsection 4.2.1.

This completes the third step of the argument advanced here: As I have indicated in section 3.2, the different uses of COMMUNICATION are patch-like, but a closer analysis of the relations between them, conducted in the current section, indicates that those connections extend beyond local relationships between individual uses, and instead form an actual *core meaning* of the concept, captured by the proposed defintion-schema, interpreted causally, and with the fourth, functional component added. Hence, despite assuming that COMMUNICATION is polysemous, I have arrived at a contradictory conclusion: namely that it is a single concept, i.e., the term is used consistently across the different spatiotemporal scales of biological organization, preserving the hypothesized scale-free structure. In what follows I explore the epistemic benefits that come from this perspective (in section 3.4) and investigate the components postulated by the definition-schema in greater detail, to make sure that the proposed account accurately captures the use of COMMUNICATION (section 3.5).

3.4 COMMUNICATION AND SIGNALLING

As indicated in the results from the corpus and mentioned above, there appears to be a pattern where the notion of "signalling" is preferred where discussing apparently communicative structures when they occur at the scale of cells or tissues. The question is—is this a specialized term associated with a particular domain but with ultimately the same core meaning, or is it a distinct concept—perhaps, after all, a properly separate patch which would ultimately lead us to rejecting the advanced argument for the unity of COMMUNICATION and accepting its polysemous nature? I consider this potential counterexample at length here, to highlight how the proposed definition schema fares in a detailed analysis of uses of communicative notions in the literature and what epistemic benefits it brings.

The introduction of the distinction between "communication" and "signalling" can be partially explained by the pervasive training of biological researchers to follow Morgan's canon. Stemming from the early days of comparative psychology, the canon, also known as the principle of parsimony (see Morgan 1894, p. 59; this has been variously criticized, e.g. Pisula 2016) states that "in no case is an animal activity to be interpreted as the outcome of the exercise of a higher psychical faculty, if it can be fairly interpreted as the outcome of the exercise of one which stands lower in the psychological scale." It connects more broadly to the rejection of teleological language in life sciences (see Mayr 1992). From this perspective—communication, which, as argued above, initially refers to phenomena related to human linguistic interactions (see section 3.3), will be a term to avoid in the context of lower scale biological processes. This is further strengthened by the analysis of the notion of "information" discussed in the Introduction, and continued in chapter 4. In that debate, a number of scentists have opted for some time to interpret that concept in semantic terms (e.g., Maynard Smith 2000)—in that context attributing "meaning" to genes or low-level biological mechanisms (such as Notch signalling) is seen as unnecessarily "anthropomorphising" the processes.

Such arguments are raised not only in the context of communication. Consider the term "immune memory." In a recent attempt to develop a systematic strategy for assessing allegedly mnemonic systems, David Colaço (2025) analyses that concept, focusing specifically on some of the arguments that the alleged immune mnemonic system isn't *really* MEMORY. One of those is the "mere causal" argument. This argument states—in line with the spirit of Morgan's canon—that "unlike conventional memory, immunological phenomena can be described and explained in causal terms alone", hence it is not *really* memory (Colaço 2025, p. 5). This argument does not reject the causal underpinnings of "conventional" memory. Instead, it states that for those processes an additional level of description or explanation is required, as a merely causal description can't capture the complexities of "regular" memory—given the cognitive limitations of human scientists. On the other hand, proponents of the "merely causal" view (e.g., Melander

1993) suggest that immunological processes are simpler and can eventually be captured with only a causal vocabulary. A parallel reasoning likely motivates the preference towards "signalling" over "communication" at the scale of cellular and subcellular processes.

Alternatively, we can view this use of "signalling" through the work of Daniel Dennett, who coined the term "intentional stance" (Dennett 1987; developed from Dennett [1969] 2010; Dennett 1971). Dennett's notion captures a particular "strategy of prediction and explanation that attributes beliefs, desires, and other 'intentional' states to systems – living and nonliving – and predicts future behaviour from what it would be rational for an agent to do, given those beliefs and desires" (Dennett 1988, p. 495). Crucially, the use of intentional stance is pragmatic: we can apply it to the various systems because it offers us a good practical (explanatory, predictive) grasp on their behaviour, better than other possible ("physical" or "design") "stances." It does not require consideration whether the system is *really* intentional, whatever we take it to mean. This notion has been extended to other areas—for instance, Stephen Cowley proposed a "language stance" to account for the human ability to view verbal activity as language (Cowley 2011). Similarly in the current context, we could consider a "communication stance," a strategy of interpreting system's behaviour in terms of communication (rather than "merely causal" interactions), justified by the novel predictions and explanations it enables.

If we look at the most characteristic collocations for "signalling" and "communication" as identified by Word Sketch, we see indeed that signalling is most strongly associated with the cellular and molecular processes: its most significant modifiers include: "wnt", "catenin", "notch", "hedgehog", or "bmp", which are names of particular signalling molecules or molecular pathways involved in development; as well as more straightforward references to scale: "cellular", "intracellular" or "intercellular." Nouns modified by "signalling" include "pathway", "cascade", "technology", "process" or "capacity." It is important to note that cell-related modifiers are shared between "signalling" and "communication", as well as the three latter nouns modified by the two terms. Similarly, the verbs which take the terms as object are either the same—"mediate", "enhance"—or semantically related, e.g., "activate", "induce" or "initiate" for "signalling" vs. "enable" or "establish" for "communication." Terms related to signalling have more "technical" collocations—with "induce", "transduce" or "antagonize" being good examples. However, overall they seem to fit in similar categories—with references to the mark being transferred (particular molecule types), the entities involved (cells), and the structure underlying the process (identified as a pathway or a cascade).

Alternatively, we can look at topic similarity for both terms[9]. Specifically, the five topics that have the closest similarity scores for both "communication" and "signalling" (i.e., the smallest difference between the scores for both words) are topics WESTERN BLOTTING (25)[10], PLANT SYMBIONTS (24), GUT MICROBIOME (17), ENERGY METABOLISM (40), and FUNGAL INFECTIONS (13). These topics have all relatively low topic similarity scores to both search terms (these are in the range of 0.286 – 0.359, with the means for "communication" and "signalling" across all topics equal to 0.445 and 0.393, respectively). Hence, the similarity is rather uninteresting. It is more interesting to look at the largest differences between similarities to the two search terms. The 5 topics that have the most different similarity scores have high similarity to one of the terms (4 out of 5 to "communication") and an average one for the other. These topics are: SOCIAL FACTORS IN LANGUAGE (5, 0.834 for "communication" and 0.365 for "signalling"), GENETIC EVOLUTION (27, 0.755 and 0.374), NOTCH DEVELOPMENT (6, 0.444 and 0.792; the only one more similar to "signalling" than to "communication"), COMMUNICATION IN EDUCATION (14, 0.699 and 0.357), SOCIAL MEDIA (26, 0.680 and 0.348). NOTCH DEVELOPMENT is a topic that refers to a set of molecular pathways involved in development[11]. Hence, it is used in this context as a highly specialized term, referring to a particular protein-receptor pair. The only other topic that is strongly associated with "signalling" is QUORUM SENSING, which is, however, also relatively similar to "communication" (0.519 for "communication", 0.705 for "signalling"). This topic refers to low-level processes, but "signalling" and "signal" are, in this context, more often used as adjectives associated with "molecules", "receptors", or "response system." The next three topics have much lower similarity scores, but exhibit a similar pattern: ION CHANNELS (12, 0.497 for "communication" and 0.572 for "signalling"), FLUORESCENCE IMAGING (28, 0.528 and 0.558) and MECHANICAL SIGNALS (10, 0.370 and 0.530). ION CHANNELS refers to lower-level processes. FLUORESCENCE IMAGING is in fact a methodological use of the word "signal", as it refers to the signals that researchers identify in

---

[9] As BERTopic operates on embeddings, it can represent both the topics and individual words or phrases in the same vector space, and compare the (cosine) similarity. Note that the analysed subcorpus has been generated using several definitions of "signal" (queries 19-28, as discussed in subsection 3.2.1), hence there is no a priori reason to believe that the topic-similarity distribution should be skewed towards "communication."

[10] Western blotting is a technique for detecting specific proteins in tissue sample, widely used in molecular biology (see Kurien and Scofield 2006). It involves the use of specific synthetic or animal-derived antibodies: a primary, which bind to the proteins and a secondary, which binds to this initial antibody and can be later imaged using immunofluorescence or staining. This topic is in fact an artifact—it's been picked up by semantic search due to the fact that one of the companies that produces antibodies used is called "Cell Signalling Technology, Inc." Due to the standardised manner of reporting the methods used for western blotting, these paragraphs have been grouped together by topic modelling. However, this is not an entirely irrelevant context here, as these methods of analysis of biological material is an ingenuous use of cell's *in vivo* signalling mechanisms.

[11] See a more detailed discussion in footnote 5 above.

their analysed material (e.g., includes the notion of "cross-signal" which refers to the artifacts arising from insufficient specificity of molecules used for the analysis). Finally, MECHANICAL SIGNALS involves mainly the adjective use of "signalling", concerned with the biochemical details of cells exerting and receiving influence on one another via mechanical forces.

What this suggests, is that there is a particular domain that is more strongly associated with "signalling." In the context of low-level biochemical processes this term has a more specialized meaning. However, even in this context, "signalling" is used rather to describe the *mechanism* of biological communication. As seen in the Word Sketch results above, it is differentiated primarily by the particular material implementation of the communication process at that scale, while preserving the key components postulated in the definition-schema. This technical use is likely motivated by the deflationary character of "signalling", as compared to "communication." This seems, however, not to be a highly systematic distinction, and indeed one that largely follows the patterns of use of the two terms in natural language. Consider the following scenario: during a hike in remote mountains we see a human figure on the path, a few kilometres distant. The person is waving and gesturing unclearly, as they do not seem to be in any distress. In these circumstances, it would be perfectly admissible for me (if a bit artificial), seeing this person, to say to my hiking partners: "Look at that person, they're *signalling* something! What do you think they're trying to *communicate*?" One systematic difference that is visible here is that "communication" in this context seems to presume the reception of the message, hence, its success. "Signalling" does not have such a presumption—the success criteria seem to be distinct, and just sending the message leads to an instance of successful signalling. Note as well, that in this context both terms—in fact—admit the involvement of semantics. The distinction related to the success criteria is, however, associated with the distinct ways that semantics are characterized across the scales of biological organization—as I will discuss in chapter 5—and does not overwrite the shared core meaning across "signalling" and "communication."

3.5 COMMUNICATION, NATURAL SIGNS, AND CAUSALITY

Before concluding the analysis of COMMUNICATION, I would like to consider an important objection. It is a justifiable consideration, especially for philosophers reading this, to look at what I have suggested as the core meaning of COMMUNICATION and wonder whether conditions put forward in the definition-schema are necessary or sufficient. Indeed, it seems that there are several biological (and perhaps non-biological) processes that apparently meet the outlined criteria but would not normally be classified as cases of BIOLOGICAL COMMUNICATION. The definition-schema of COMMUNICATION proposed does not indicate whether these should be regarded as necessary or sufficient conditions. Henc, we need to consider some of the examples that are apparently included in my proposed account, but the actual use of the term would indicate otherwise. To highlight the scope of what I'm proposing here, consider the following cases:

1. A billiard ball rolls down the table and hits another. We have a transfer of a mark here (following the standard *mark method* to which I purposefully referred in the formulation of the conditions, see Reichenbach (1958) and Salmon (1984, 1994), where mark is understood as "a local modification in structure", see Salmon (1984, p. 147)—in this case, the kinetic energy and the associated possibility of one ball indenting or breaking the other, for instance), we have separable entities in the form of the two balls, we have the structure provided by the table. But it's not communication!

2. Bacteria in the centre of a growing biofilm start to deplete the nutrients and the amount of ammonium they produce and diffuse throughout the biofilm starts to decrease. This results in the halting of the cell division of the bacteria on the periphery, slowing down the growth of the colony until the nutrients, specifically glutamate used for the production of ammonium, available on the outside can diffuse to the centre (see Liu et al. 2015). There is a mark (ammonium) being transferred, or rather—withheld, there are separable entities (individual bacteria), and a structure (the biofilm medium) that enables this process. Is it communication?

3. Bacteria in the centre of a growing biofilm start to deplete the nutrients and in response depolarize by releasing potassium, which is actively propagated throughout the biofilm, until it reaches the periphery. The depolarization wave limits the ability of those cells to take up and retain nutrients, allowing them to diffuse through the biofilm and reach the cells in the centre, effectively synchronizing the metabolic states of cells across the biofilm (see Prindle et al. 2015). Similarly, all criteria for communication set here (potassium as the mark, individual bacteria as entities, biofilm medium as the structure) are met and further this process is widely described as "signalling" - is this communication?

4. The digestive system breaks down carbohydrates into monosaccharides, such as glucose, which get released into the bloodstream and stimulate the pancreas to secrete insulin. We have a transfer of a mark here, in the form of glucose, we have separate entities—the digestive system and the pancreas, we have a structure—the bloodstream, the specific receptors that pick out concentration of sugars. But it's not communication!

5. A predator picks up the scent of a prey. We have the transfer of a mark, the pheromones left by the prey animal, e.g., on their faeces, we have distinct entities involved and a structure which evolved for this particular function. Yet again, it's not communication![12]

Examples with billiard balls (Case 1) and ammonium in biofilms (Case 2) are intended as two different examples of "merely causal" processes, with the example of potassium signalling in biofilms (Case 3) being intended as a contrast class. Examples with glucose (Case 4) and scent (Case 5) are instances of natural signs, and again—the example of potassium signalling (Case 3) will serve as a contrast class. I discuss them in turn in the remainder of this chapter. While the problem of distinguishing signals from cues—or natural signs from instances of communication—has received some attention already, the section that discusses the causal constraints on communicative processes offers an answer to a question which has hitherto escaped philosophical scrutiny. This analysis allows me to further specify the components postulated in the definition-schema.

3.5.1 *Communication and causality*

One widely accepted theoretical virtue in science is that of simplicity (e.g., Longino 1995): if two explanations of a phenomenon provide precisely the same epistemic benefits, the one that makes fewer assumptions and is overall simpler will likely be preferred by scientists. Ultimately this seems to be the motivating factor for accepting the Morgan's canon (as discussed above, in section 3.4) and opting for a "merely causal" description of processes, without intentional categories (see also section 3.3 and 5.4.1), whenever possible. But besides impacting the preference towards "signalling" vs. "communication", it is also a relevant factor in considering whether to employ a communicative explanation *at all*. Indeed, in biology, there is a class of "merely causal" processes, which motivate more "mechanical" (in a broad sense) accounts, even if metaphorically can be described as communication. Using the Dennett's notion of the stance, in these cases a physical stance suffices. In contrast, communicative accounts can be more appropriately applied in cases where the putative signaler—or the cause—is *removed from immediate physical effects*. One way in which this can be achieved is through

---

[12] Note that there is a class of genuine communicational signals that prey animals send to their predators, e.g., the gazelles stotting to inform the predators that they have been spotted (see Artiga 2021)—and resultingly avoid the unnecessary expenditure of energy for both animals.

the arbitrary relation between the signals and responses they evoke.[13] Another is through a material discontinuity in the causal process. I explore these notions through the cases suggested at the beginning of this section, highlighting the functional and causal considerations emerging from the corpus data.

To highlight the intuitions behind this, let's consider the cases presented above in some detail:

Billiard balls

In the case of billiard balls (Case 1), we could try to account for what happens by saying that the first ball, let's assume it's the white cue ball, signals to the second ball, let's assume it's a green striped ball, to move. The cue ball is the sender, the green ball the receiver. The signal—the mark transferred—is the kinetic energy the ball transfers to the other. The structure that enables this is the surface of the table, which constrains the dimensions of the balls movement, and further the vector of the cue ball's movement, on which the success of this instance of 'communication' hinges.

But this description is quite absurd and doesn't provide any benefits compared to the purely physical account in terms of the transfer of kinetic energy. Besides the implicit suggestion that the ball somehow performs the action of the signalling and ascribing some agential properties to it this way, the absurdity of this description hinges also on the fact that the response evoked by the putative signal is an immediate physical effect. Nevertheless, even if a description in terms of "communication" doesn't really apply in this instance, there are accounts of causality in terms of information flow or transfer—one such account has been suggested by John Collier (e.g., Collier 1999, 2011). I come back to this in subsection 3.5.2 below.

Biofilms and ammonium

In the case of the biofilm (Case 2), specifically *B. subtilis* biofilm, and ammonium, we could provide a story in which bacteria at the centre of the biofilm stop producing ammonium in order to signal to the periphery that they have depleted nutrients, glutamate specifically, and the bacteria on the periphery of the colony notice the absence of ammonium and in response slow down cell division, allowing the glutamate to diffuse. We have here the hallmarks of communication identified before: separable entities, in the form of central and peripheral cells, a mark that is being transmitted, in this case in the characteristic form of change, showing that signals need not correspond to specific physical entities, but also to patterns. Finally, we have a structure in the form of both the medium in which the biofilm is embedded, and which, for example, impacts the dynamics of the

---

[13] I owe the recognition of these properties of communication to the discussion with the members of the Levin Lab at Tufts University, in particular with Franz Kuchling and Santosh Manicka, where I presented some of this work in August 2022.

diffusion of ammonium and glutamate, and in the form of the dynamics of the ammonium production before it is switched off, which allows the change to be detectable by peripheral bacteria.

But this description is not true (the following account is based on Liu et al. 2015): ammonium is a substrate required for producing glutamine (essential for growth) in a process where an enzyme, glutamine synthetase, combines it with glutamate (Gunka and Commichau 2012). Ammonium can be produced by the bacterium itself from glutamate, but as ammonia vapor, which is in equilibrium with ammonium, can cross cell membrane and disperse, this process is known as a "futile cycle." This "futility" is further enhanced by media flow which is higher on the periphery of the biofilm. Resultingly, bacteria prefer to absorb extracellular ammonium and this preference will be stronger at the periphery than in the center, where the flow is smaller and in result the loss of ammounium is smaller as well. In effect, the biofilm metabolism can be modeled by considering the centre of the colony as producing ammonium for the periphery. This process, however, depends on the availability of glutamate: the depletion of glutamate at the centre halts the production of ammonium, which in result makes it impossible for the peripheral organisms to produce glutamine and stops their growth. This is a strictly causal connection between the metabolic cycles of bacteria in the biofilm, which leads to a co-dependence of individuals within a colony (and enhances the adaptive character of this organization of bacterial life, which is more likely to occur under environmental stress, see e.g., Berk et al. 2012). The change in the production of ammonia in the centre of the biofilm is an immediate effect of the depletion of glutamate, and the lack of ammonium immediately causes the halting of growth of cells at the periphery. The signalling gloss, again, doesn't provide any epistemic benefits and in some ways obscures the actual causal structure of the process.

Biofilms and potassium

In the case of the *B. subtilis* biofilm and potassium (Case 3)), the communicative description is the one preferred by researchers (see Prindle et al. (2015), on which the account here is based; note that this process occurs in parallel to the ammonium one just described). When the bacteria at the centre of the biofilm deplete the available glutamate, the change in the metabolic state of the cell triggers potassium flux through the ion channel (Prindle et al. (2015) investigate the YugO channel, which is gated by a TrkA protein known to change conformation in response to ATP binding—a process susceptible to the metabolic state of the cell, see Y. Cao et al. (2013)), resulting in a depolarization of the cell. The increase in extracellular potassium further triggers a short-term membrane depolarization in neighbouring cells (at least partially through the YugO channel, as established by investigating a *yugO* deletion strain of *B. subtilis*; note the different capitalizations for the name of the protein and the name of the gene), which actively propagates the potassium—and depolarization—wave throughout the biofilm. As this wave spreads, through the "link between membrane potential and metabolic activity" (Prindle et al. 2015, p. 5; this happens through physical interactions—as glutamate and ammonium are charged molecules, the change in the voltage of the cell membrane impacts the electrical forces mediating interactions between the cell and the nutrients. The relationship is quite recognized, see e.g., Hulbert and Else 1999), it limits the ability of peripheral cells to take up or retain nutrients, altering the underlying metabolic reactions, but not contributing to them. In result, the growth slows down and the nutrients can diffuse to the centre of the biofilm. When considered through the communication criteria identified, we again see that the process involves separate entities—central and peripheral cells, a transfer of a mark, in the form of the potassium wave, and a structure that involves the active propagation of the depolarization wave by intermediate cells and the causal links between metabolic state and signalling triggers, and, inversely, reception of the potassium signal and the metabolic activity.

While there are undeniable causal connections between the depolarization and metabolism, the signalling gloss provides a certain simplicity by indicating a separate mechanism which is not metabolic in nature, i.e., doesn't directly produce energy, but rather takes up some of it, and in turn affects metabolic processes. This benefit seems to go beyond the fact that the signalling gloss provides natural stand-ins for some of the components of this process that are still lacking a molecularized mechanical description, but rather hinge on the fact that it forms a parallel, non-metabolic channel for synchronizing metabolic states across the biofilm. This channel further provides important evolutionary benefits, as the potassium spread is faster, allowing for rapid response to nutrient depletion and ensuring the health of the biofilm. It is also likely not limited to the individual colony's bounds, and in principle could metabolically synchronize physically disconnected neighbouring biofilms (see Prindle et al. 2015, p. 5).

Considering these examples, the limits of the proposed criteria for communication become clear: while both examples involving biofilms (putting the case with billiard balls, (Case 1) aside for the moment) meet the laid out criteria, only the example in subsection 3.5.1, with potassium ion flux, is explicitly discussed in terms of signalling or communication. I will argue that the difference between those two cases boils down to the character of causal links within the structure of these processes so that ion fluxes count as signals and ammonia fluctuations don't, because of the *removal from immediate physical effects*. While this category—in my view—phrases the intuition behind this distinction nicely, it requires some philosphical work to tie it directly to existing concepts, such as that of "detachment" (see Moss 2009). For this reason some work needs to be done to highlight what the "removal" means in this case and in what way introducing this distinction aids us in the causal analysis of biological communication.

Causal analysis of biofilm interactions

To lay out the argument, let's consider the two examples of the centre halting the growth of the periphery in response to depleted nutrients (Cases 2 and 3). I will use here the standard approach of depicting causal structure in the form of causal graphs (see Pearl 2022; Woodward 2004). I will not be able to offer a full causal model, as this would go well beyond the scope of the thesis, but instead will focus on the structure of the processes analysed. To this end, I will follow a broadly interventionist view of causation (in the sense of Woodward 2004, 2023), the dominant view of causation in philosophy of science.

Broadly speaking, the interventionist theory of causation is a species of manipulability accounts of causal processes, which view causes as "handles or devices for manipulating effects" (Woodward 2023). Interventionists hold that "X is a cause of Y iff there is a possible intervention on X that changes Y" (Reutlinger 2013, p. 25), where an intervention is understood as a primitive counterfactual notion. An intervention on a causal variable X is an operation which sets the value of that variable, disrupting the relationship between X and its parents (in the causal graph) and leaving all other causal relations unchanged (as it is counterfactual, it does not need to be an *actually* possible intervention). This approach offers the most suitable tools for the foregoing argument, as it fits well with the scientific methodology, where experimental manipulations are used to investigate causal mechanisms.

I will draw here heavily on the work of Lauren N. Ross, who has been consistently applying the interventionist framework to biological and medical contexts. However, as manipulationist theories of causation are an active area of research, some of the details might be susceptible to different interpretations, depending on the exact interventionist theory accepted. I hope that this variety will not detract from the main line of reasoning as the argument here does not—in my view—hinge on the details of the account of causation.

As I'm considering type-level biological causes, the resulting graph involves some cycles, a common feature of biological causation (see Bechtel and Bollhagen 2024; Gebharter and Leuridan 2024), which complicate the structure (and more practically, the calculations of probabilities involved, though this is beyond the scope of this thesis)—however, these cycles are not vicious in any way, and if we were to consider a token-level description of the same processes, they would unravel (in a time-indexed form). Finally, for the purpose of simplicity, the analysis will be done considering the variables, properties and entities involved, even though I believe that a more correct depiction of this should involve a processual account.[14]

---

[14] In fact a processual interventionist account of causation has been jarringly overlooked in the rapidly growing literature on process ontology, which has opted for accounting for causality in other ways (e.g., through an account of dispositions, see Anjum and Mumford (2018)), but does not explicitly argue against interventionist accounts either as an ontology of causation or as means of causal discovery. One exception in this regard is the work of Peter Machamer—the new mechanist approach he developed (Machamer, Darden, and Craver 2000) was motivated by

The causal process depicted in the graph 3.5 combines the processes described in Section Biofilms and ammonium, presented in the top row, and in Section Biofilms and potassium, presented in the bottom row, as they share a starting point (glutamate concentration in the centre of the biofilm) and the destination (the growth of cells in the periphery). The senders and receivers (or causal agents) are not included explicitly as the descriptions in Liu et al. (2015) and in Prindle et al. (2015) do not presuppose agency on the part of the entities involved, but rather consider directly the individual events that constitute the causal chain: changes in observable properties of the entities (bacteria and the extracellular environment). Both causal chains are traditionally considered as causal pathways (in the sense of Thagard (2003); Ross (2018); Ross (2021a)) controlling the growth of the cells at the periphery (considering the criteria in Ross (2018), p. 557): they are sequences of causally connected events that offer (i) fine-grained causal control over the variable of periphery growth (as the rate of growth depends on the concentration of the ammonium or the charge of the extracellular environment in a continuous manner, see Liu et al. (2015), Extended Data Figure 1), that rely on (ii) material continuity (see below; notably later Ross eases this condition by stating instead that pathways should involve "the flow of some entity or signal through a system", Ross (2021c), p. 14), and that (iii) have a fixed order of steps shared across some large domain of systems (in the case of the metabolic chain it is likely limited to colonial unicellular organisms, but the signalling chain is conjectured to expand over a broad range of systems as Prindle et al. (2015), p. 6 consider that "it is intriguing to think not only about the structural similarities between bacterial and human potassium ion channels, but also their possible functional similarities with respect to long-range electrical communication").

The character of causal links still seems to be a better guide in drawing the distinction we're after here than the causal structure of the process itself. Indeed, both pathways can be considered as providing the contributing causes (in the sense of Woodward 2004, ch. 2) for the growth of the cells at periphery. Consider the connecting point of the two pathways: both ammonium concentration at the periphery and the depolarization wave are (putative) direct causes of `Gln` production with respect to one another. If an intervention is performed on the amount of ammonium in the medium (as reported in Liu et al. 2015, Extended Data Figure 3; I_4 on the graph) without influencing the membrane voltages, it results in a change of the biofilm growth dynamics (the oscillations in growth rate are suppressed). Similarly, the experiments in Prindle et al. (2015) highlight that an intervention into the extracellular potassium concentration (I_5 on the

---

the interventionist account, and Machamer himself (although that has not become a dominant view in this line of thinking) adopted a process ontology (e.g., Machamer and Springle 2016). While I do believe that there is no tension between processual ontology and interventionism (I think their succesful merger falls out of Karen Barad's agential realism, see Barad (2007), and chapter 2; this is likely also a view compatible with the classical Whiteheadian views on causal efficacy presented in Whitehead ([1927] 1985)), I do not have the space here to discuss this in any detail and for this reason I will stick to the broadly accepted form of description through "static" variables and properties.

Figure 3.5: (see next page) Causal graph for biofilm metabolic co-dependence and potassium signalling processes. GLU stands for glutamate, Gln for glutamine, NH_4^+ is ammonium and K^+ is potassium ions. Nodes connected with wavy arrows depict the interventions introduced in (Prindle et al. 2015; Liu et al. 2015), on which the graph developed here is based on. Their interventions included the addition of GLU, Gln, NH_4^+, GDH (glutamate dehydrogenase, enzyme necessary for the production of glutamine from glutamate and ammonium), and KCl to the growth medium, as well as the use of genetically modified strain of bacteria with the gene coding for the TrkA protein deleted, effectively rendering the YugO channel inoperant. The boxes depicted in the bottom row are introduced to highlight how the description of the process in (Prindle et al. 2015) switches between explanatory levels. E.g., the difference between potassium flux and the accompanying depolarization wave, while highlighted as if these where separate processes or properties, amounts to the different aspects of one process being highlighted: either the change in extracellular potassium concentration or the membrane voltage; while the latter is constituted by the former, the assumed impact on metabolism occurs “through” the change in membrane voltage, rather than extracellular potassium concentration, while the propagation of the wave is effected by the potassium concentrations. Further, the depicted cycles could be disambiguated if we were to consider a token-level causal process, but this would result in a more confusing diagram. Note that the diagram is connecting both mechanistic and phenomenal properties of the process. As it was meant as an illustration, readability was prioritized; it would be more accurate to depict both processes entirely mechanistically, as components and their activities, without relying on observable properties as is done here. This however would hinder the clarity of the graph.

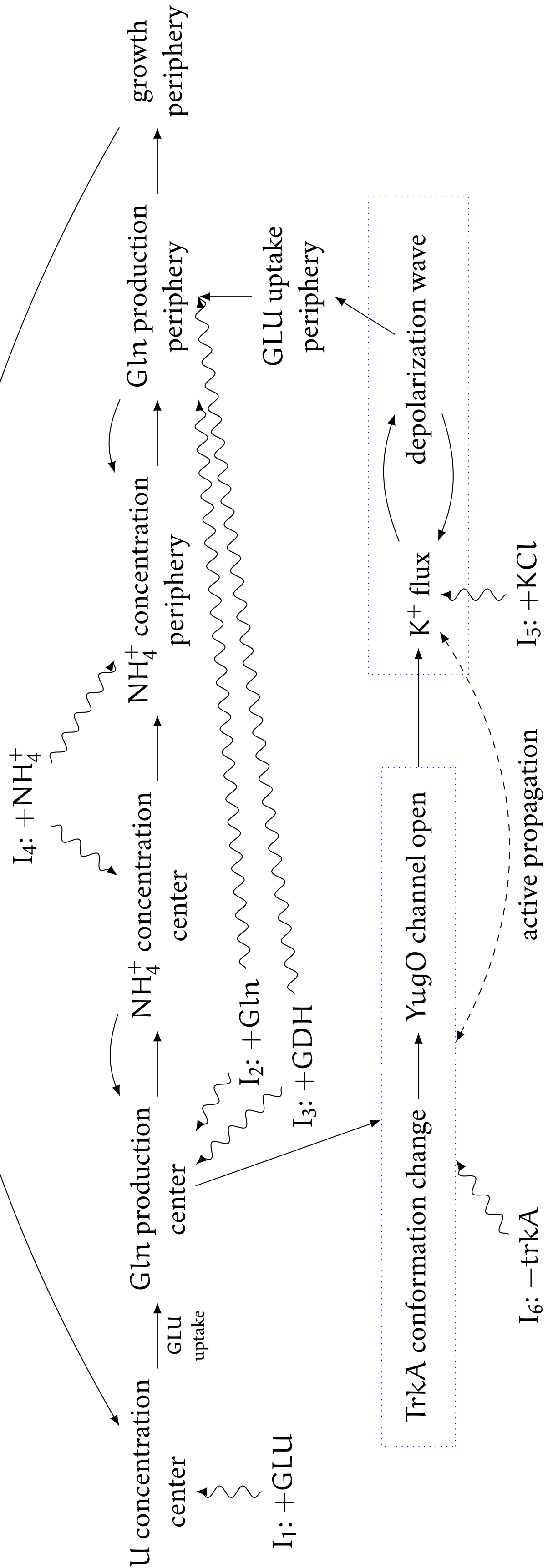

GLU concentration center
I_1: +GLU
GLU uptake
Gln production center
I_2: +Gln
I_3: +GDH
NH_4^+ concentration center
I_4: $+NH_4^+$
NH_4^+ concentration periphery
Gln production periphery
growth periphery
GLU uptake periphery
TrkA conformation change
YugO channel open
K^+ flux
depolarization wave
active propagation
I_5: +KCl
I_6: −trkA

Figure 3.6: Causal graph depicting abstracted structure of biofilm growth.

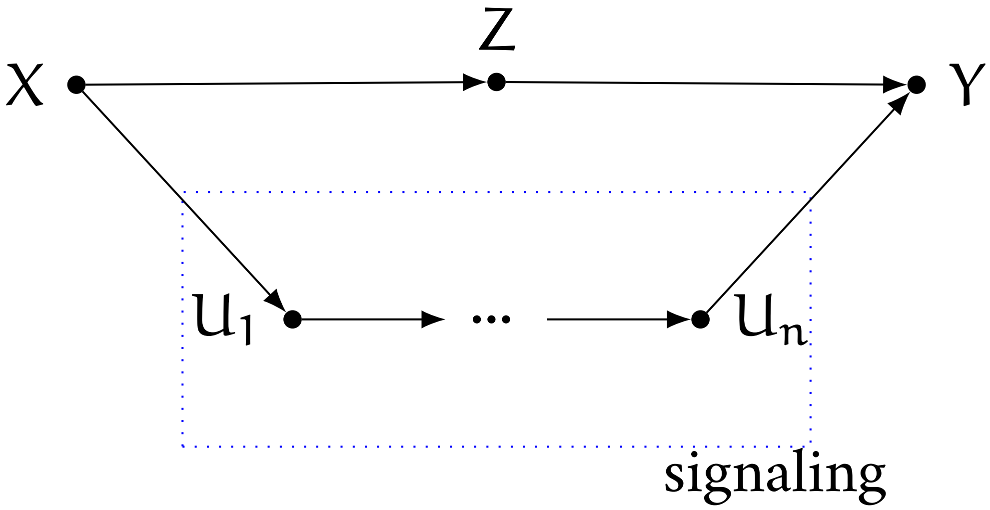
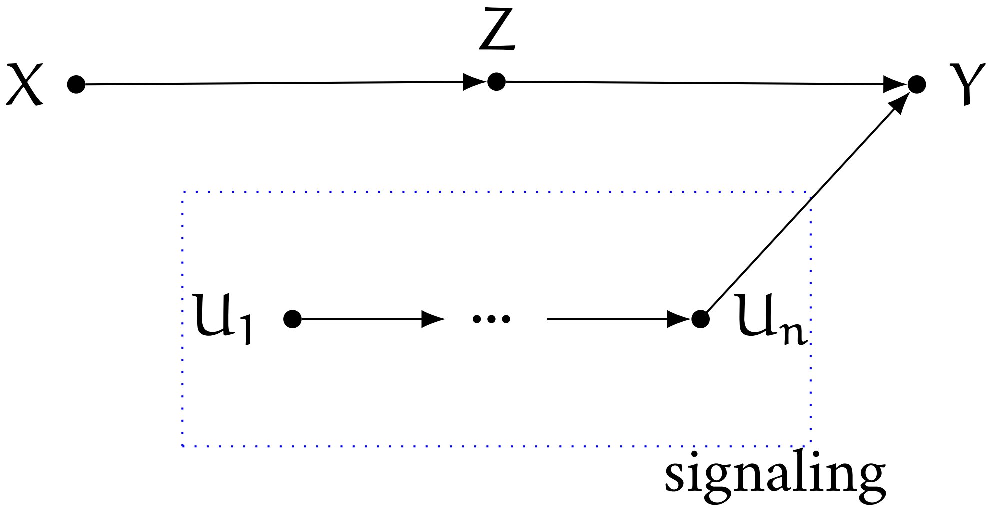


Figure 3.7: Causal graph depicting abstracted comparison of causal structures between "merely causal" and communicative processes with shared effect.



graph) impacts the oscillations of membrane voltage.[15] They also directly test the link between glutamine production and depolarization by supplying the biofilms with glutamine in the medium (I_2 on the graph).[16]

To account for the difference at play here, I suggest to have a closer look at the criterion of "material continuity" (expanded on in Ross 2021b). This is a particular version of the condition of the transfer of the mark that I have suggested (when proposing a tentative communicational account of the processes). In the two cases analysed here, it corresponds to the propagating changes in ammonium concentration or the potassium ion flux. These examples quite directly meet the criteria laid out by Ross: the carry-over of material is significant, reliable, and constrained to the causal process (see Ross 2021b, p. 6).

Ammonium in the metabolic chain is the direct substrate of the reactions at later steps of the pathway and as such the physical material gets preserved (analogous to the analysis of glycolysis in Ross 2018, pp. 553-555). The signalling chain is slightly more interesting, as Ross explicitly considers causality through signalling as a contrast class to causes with material continuity (2021b, p. 13), even in the cases where the two are associated and can be studied through one another (Ross 2021c, pp. 3-4). She analyses material continuity in the context of tracer techniques - research procedures widely used across life and social sciences which consist in marking entities and tracing the trajectory of the mark. One of the most widespread examples in biology is marking substances with radioactive materials to track the subsequent events they are a part of—biochemical reactions, cells they interact with etc. Indeed, Prindle et al. (2015) do employ tracer techniques to study the role of potassium fluxes. They use voltage sensitive dyes Thioflavin T (ThT) and Asante Potassium Green-4 (APG-4). ThT is positively charged and as such can be retained by cells which have negative membrane potential, resulting in serving as a faithful reporter of the membrane potential. APG-4 (of the two, this dye is more appropriately to be considered a "tracer") is used to study the extracellular potassium concentration as it increases fluorescence in response to the potassium cations concentration. This allowed Prindle et al. (2015) to trace the spreading potassium wave through the biofilm (Figure 2). While this kind of continuity is more statistical,[17] as it does not depend

---

[15] This is strongly correlated with the growth rate, due to the fact that both ammonium and glutamate are electrically charged metabolites, "whose respective uptake and retention is known to depend on the transmembrane electrical potential and proton motive force" (Prindle et al. 2015, p. 2). This in a sense "delegates" the explanation of the mechanism in action to physical properties of electrically charged entities.

[16] Supplying extracellular glutamine omits the need for the production of the amino acid and allows continuous growth to the bacteria at the periphery, as shown in Liu et al. (2015), Extended Data Figure 3. Further, it stops the membrane voltage oscillations, see Prindle et al. (2015), Extended Data Figure 1.

[17] Compared to the paradigm case of the role of an egg in the development which Ross considers (see Ross 2021b, p. 5). In that case "The substrate at each upstream step is causally relevant to the creation of the most immediately downstream product and there is material that reliably flows along this causal process. In other words, the same material that makes up the cause, becomes part of the downstream effect" (Ross 2021b, pp. 5-6).

on the *individual* molecules but on their concentrations (in a similar way as in the ammonium example), it *is* a case of material continuity: e.g., it is not true that the electrical charge of molecules is the difference-maker, due to the mechanistic dependence of this process on potassium channels in active propagation of the signal (as ascertained by the intervention I_6 depicted on Figure 3.5, see Prindle et al. 2015).

While clearly the presence of material continuity does not serve as a clear indicator of "removal from immediate physical effects", it is important to note at this point that in the more strict version of Ross' formulation, this criterion also needs not to be met by biological signalling. To remain within a familiar realm, consider the role of glutamate as the major excitatory neurotransmitter in the nervous system. Throughout the various cells present in the nervous systems, glutamate is involved in signalling either via *receptors* (such as the N-methyl D-aspartate (NMDA), α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA), and kainic acid, see Pankevich et al. 2011) located post-synaptically, or via *transporters*, which are located either pre-synaptically, or post-synaptically e.g., in astrocytes (see M. B. Robinson, Lee, and DaSilva 2020). While the transporter-based glutamate signalling apparently meets the criteria of material continuity, the receptor-based version seems to be a case of "signal continuity." The mechanism consists in the glutamate ligand binding to the receptor and changing its conformation, leading downstream to further changes that do not involve glutamate itself (but rather ion fluxes, in the case of ionotropic receptors, or gene expression and protein synthesis, in the case of metabotropic receptors, see Pankevich et al. 2011).

This is, however, only a part of the story. If we pay attention to the cases of potassium flux in the biofilm and glutamate signalling via transporters what we note is that while both of these examples involve some material continuity, as I have argued just now, but they actually also involve breaks in material continuity. In the case of potassium fluxes, we observe them in two places: at the very beginning of the signalling pathway, where the binding of ATP to the channel protein leads to the depolarization and potassium outflow, and at the very end of the pathway, where the potassium depolarizes the membrane of cells at the periphery preventing them from taking up glutamate from the medium. Similarly, if we consider glutamate signalling in astrocytes: glutamate is transported, ensuring material continuity between the presynaptic neuron and the astrocyte, but then the continuity is broken as the signal is transduced into the cell depolarization (glutamate transporters are dependent on sodium cations and the rapid intake of glutamate results in a significant depolarization, which leads to the repolarization and the associated movement of ions—"contributing to signalling within and between astrocytes", see M. B. Robinson, Lee, and DaSilva 2020).

Hence, I would like to suggest that the "removal from immediate physical effects" is to be best understood as the presence of material discontinuities. In particular, these discontinuities will most likely be implemented with *activities* (in the sense of new mechanism, see Machamer, Darden, and Craver (2000);

Bechtel and Bollhagen (2021)) as their interface: the presence of large number of potassium ions depolarizes the membrane which inhibits the *glutamine uptake* in the case of biofilms; transport of glutamate *depolarizes the membrane* resulting in the outflow of sodium cations in the astrocytes. The decrease in the ammonium concentration slows the rate of the metabolic reaction that takes ammonium and glutamate as substrates: and the continuity remains unbroken.

Importantly, the discontinuity must occur at the direct cause of the process that we consider the reception of the signal, regardless of whether the discontinuities recur at every step of the signalling pathway or happen exactly once. If we consider a tentative signalling pathway that consists of material discontinuous steps to a certain point and materially continuous steps from that point onwards, I'd argue we would consider the step marking the division as the point at which the reception of the signal occurs and the downward causal processes as the response to the signal.

One special case of this requirement of material discontinuity of the direct cause involves the arbitrary connection between the signal reception and the resulting response. Arbitrariness as a feature of language has been first characterized by Ferdinand de Saussure (see Jakobson 1980) and since has been either regarded as one of the key distinguishing features of linguistic systems, or heavily criticized. The use of "arbitrariness" I opt for takes into account the generative entrenchment (see Schank and Wimsatt 1986) of both language and biology and is meant as a counterfactual statement: while we cannot directly shuffle the signal-response pairs in biology or form-meaning pairs in language, *they could have been different*. This of course does not mean that there are no pressures for them to be as they are, either due to chemical energetics, some historical dependencies, selection pressures, etc. To use Jakobson's example: while we cannot say *fromage* for "cheese" in English, and *cheese* in French, it is possible that the two languages had developed in such a way that these two words would have been indeed swapped and saying *cheese* for cheese in English, and *fromage* in French would not have been allowed. In the context of biological sign systems, arbitrariness is most often considered in the context of DNA, where the mapping between codons and amino acids could have been different (this notion of arbitrariness in relation to DNA has been explored by Stegmann 2004; see also Deacon 2011; and the notion of "gratuity" coined by Monod 1971), but the use of electrochemical rather than purely chemical signalling in the nervous system is also arbitrary—a historical accident with particular features (e.g., speed of signal transfer) that enhanced its fitness and ensured evolutionary success (see Wan and Jékely 2021). This fact is indeed of central importance to the argument for establishing the field of "plant neurobiology" (see Brenner et al. 2006; Baluška and M. Levin 2016). If such arbitrary connections were materially continuous they could not have been regarded as arbitrary: if the response depends on the particular chemicals in the signal (as is the case in metabolic pathways, including in the case of ammonium discussed above), then the putative "signal-response" connection *could not be otherwise*.

Another way to characterize this distinction is to consider the process of receiving a signal as "detached" from other components of the communicative process. The notion of detachment I draw on here "is a measure of the relative independence of an entity from a larger milieu—its ability to resist the forces of thermal or other kinds of winds" (Moss 2009), closely relating to the notion of "internal degrees of freedom" of a given process. Lenny Moss (2009) gives an example of an atom which after being "hit by a photon, may respond to this packet of energy by translating in space, rotating, or having an electron jump into a higher energy state." This is a lower degree of detachment than that held by a biatomic molecule, which "can do all of the above plus it can vibrate along its molecular axis," and much lower than that held by larger molecules which have a broader "repertoire" of stochastic (as far as we know) reactions. In the context of a communicative system, the "thermal forces" and "winds" are understood metaphorically—detachment refers to how much the response to the signal, including its "processing", is determined by the physicalities of the particular signalling pathways. In the case of arbitrary sign-systems, such as DNA (according to Stegmann 2004), the response is clearly detached, as it is primarily causally determined by the molecular "machinery" for transcribing and translating the genetic code. In other cases of materially discontinuous processes, this causal asymmetry of sorts is retained as well, with the signal acting as a trigger to a particular causal pathway within the recipient. This feature is captured in Howard H. Pattee's account, who suggested that one of the necessary conditions for the appearance of a message is that "very complex interactions lead to a very simple result" (Pattee [1969] 2012, p. 58). While perhaps not all results of signalling are best described as "simple", the intuition that Pattee captures here is that of detachment: the response is in a smaller degree determined by the (physical) structure of the signal, effectively simplifying the process for the purpose of scientific investigation, without losing key causal characteristics.

Further, this is related to the property of signalling that Marc Artiga (2021) centers in his discussion. Artiga's approach focuses on the (teleo)semantic view of communication, as discussed at length in chapter 5, and ultimately his argument is about the content of the signal. Nevertheless, he considers similar counterexamples to that received view, as the examples discussed throughout this section. He also eventually arrives at a view which distinguishes signals due to their causal profile, although he captures the distinction by counting signals as "minimal causes." He notes the unique energetic profile of signals, as characterized by e.g., J.B.S. Haldane (as quoted in E. O. Wilson 1975, p. 176) or R. Haven Wiley (1994), according to whom signals are characterized by energetic asymmetry, where the signals themselves are energetically cheap compared to the response they elicit. Especially Wiley's (1994, p. 162) formulation, which notes that signals do not "provid[e] the power to produce the entire response" relates closely to the organizational closure and detachment of the receiving system, as discussed above (see section 1.2). Artiga defines the minimal cause as a "non-enabling cause" (Artiga 2021, p. 8587), where causation is treated in standard, interventionist terms. The notion of an "enabling" cause is less standard. Artiga considers the "ability" of a system to perform certain action, which he describes as having a mechanism (in the new mechanist sense of Machamer, Darden, and Craver 2000) for performing that action. In his analysis (a bit counterintuitively), the "ability" is a graded notion: "the more parts and activities an entity lacks for performing [an action], the less able it is to [perform that action]" (Artiga 2021, p. 8590). In that context, an enabling cause is something that makes the system able to perform the action in question. Hence, for signals, the recipient needs to be able to execute the response, regardless of the material characteristics of the signal. This notion of "minimal cause" is intended to capture what researchers mean when they characterize the role of signals as "triggering," "inducing," or "activating," picked out by Artiga and confirmed in the data discussed in the first part of this chapter.

There is, however, an important difference between Artiga's notion of "minimal causes" and my analysis in terms of "material discontinuity" of the causal process. In his analysis, Artiga aims to underscore the role of *content* of the signal. Effectively, he interprets the notion of minimal causes quite strongly, rejecting the role of *any* intrinsic (material and physical) properties of the signal vehicle. This strong interpretation is at odds with a variety of uses of the notion of "signalling" in scientific literature noted throughout the corpus analysis here. Consider only the case of potassium signals just described, where the electric charge of potassium, its physical property, is central to the ability of the molecule to act as a signal. On Artiga's account these cases would not qualify as signalling. This is a defensible position, but not an appealing one, given how the cases of potassium signalling in biofilms are investigated as evolutionary precedents of nervous systems operation, which is widely classified in terms of communication, and would likely also fail Artiga's criteria. Resultingly, I do not want to accept this conclusion, given the synthetic framing of the current research project, and the importance of the descriptive part of the work done here. I believe, instead, that the notion of material discontinuity, as fleshed out through the concept of detachment, provides a systematic way of accounting for such cases in terms of signalling, in accordance with the preferred framing in the scientific literature. Importantly, this notion does capture some intuitions present in Artiga's account—most notably, it allows for explaining the special role of arbitrariness in signalling, as already discussed, and explains the special status of sound and visual displays, which are often interpreted as signals in research (as they do not involve the transfer of matter; both these points are presented by Artiga as important consequents of his view).[18]

---

[18] There are interesting parallels between the account proposed here and the online/offline cognition distinction (see for instance M. Wilson 2002) and the concept of "representation-hungry" cognitive tasks proposed by Andy Clark (Clark and Toribio 1994; Clark 1996). Broadly speaking, Clark highlighted that depending on direct, environmental availability of information, cognition can switch between less and more representation-dependent processing. This somewhat parallels the suggestion here, with the suggested "detachment" of communication bearing similarities to the offline nature of "representation-hungry" cognitive tasks. The de-

3.5.2 *More than physics?*

Before moving on to the issue of natural signs, I need to explore in greater depth the question of why two billiard balls hitting one another are better *not* viewed as an instance of a communicative process. Clarifying this distinction will help substantiate Levy's ascription of activity to senders and receivers involved (discussed at the end of subsection 3.3.1). While I have declared the "communicative" description of the billiard ball interaction as absurd, there are in fact serious theories of causation that begin with the notion of "information."

One such account has been developed by John Collier (1999; 2011). Collier adopts a notion of information in terms of compression (see subsection 4.1.3) and claims that such notion of information necessarily has a physical form[19] and is produced by instances of symmetry breaking in physical processes (Collier 1996; Muller 2007). The resulting definition of information states that "P is a causal connection in a system from time t_0 to t_1 if and only if some particular part of its form is preserved between states s_0 and s_1 from t_0 to t_1" (Collier 1999, 2011). The notion of "form" Collier invokes is accounted for in terms of algorithmic information theory: the informational complexity of an object (an extension of Kolmogorov complexity, independent from a choice of the Turing machine—for more details, see subsection 4.1.3) measures the form of that object, though the more general intuition is the familiar notion of e.g., geometric form (see Collier 1999). As such, Collier sees the following as an equivalent statement of his principle: "P is a causal process in system S from time t_0 to t_1 iff some particular part of the information of S involved in stages of P is identical at t_0 and t_1" (Collier 1999).

The transfer of information at play is understood in terms of Barwise and Seligman's information flow (Barwise and Seligman 1997; Collier 2011, see the detailed discussion of this notion in subsection 4.1.3), and Collier adds the condition that the flow needs to be consistent with the Negentropy Principle of Information (Schrödinger 1944; Brillouin 1962), which defines the information content, which I will call NPI-content[20], of an object as the difference between

---

bate surrounding the notion of representation is alive in philosophy of mind (e.g., Smortchkova, Dołęga, and Schlicht 2020), and an important line of research seems to ground the notion in other terms, seen as more foundational for biological systems (e.g., Miłkowski 2013). In this context, this analogy could be further developed into an account grounding mental representations in biological communicative interactions. However, I do not have the space here to explore this idea further.

[19] This is an important, widely accepted claim, that is at the same time lacking a detailed consensus view. While detailed discussion of this problem goes beyond the scope of the current thesis, I mention this issue throughout chapter 4.

[20] Note that the NPI is not directly related to the notion of "information content" as Kolmogorov complexity, discussed in various places in this thesis. The issue of their relationship requires the exploration of the mentioned problem of the relation of information to its physical carriers that goes beyond the scope of this thesis—let's just note that even the Collier's account, which attempts to work out those details, does not apply well to the biological case, as it considers almost exclusively closed systems, while all living organisms are open physical systems. To avoid

maximal possible thermodynamic entropy of the system (at equilibrium) and its actual entropy: $I_P = H_m ax - H_a ct$. The Negentropy Principle of Information is closely related to the notion of Gibbs free energy, which defines the ability of a (closed) system to do useful work—which Collier interprets as the ability to control a physical process (Collier 1999). This additional condition is intended to prevent the inclusion of some spurious and indirect causal connections: "NPI ensures that if information is not lost [though it can dissipate—WR], the causal process is temporally symmetrical, and there is no internally defined temporal direction [but one can be defined e.g., through dissipation, if it occurs—WR]" (Collier 2011).

Let's consider the billiard balls example through the lens of Collier's account. The first ball rolls in a particular direction and force, and has a particular shape which constrains its interaction with the other ball. If we consider the two balls and the environment as a system, then by Collier's definition, we have information transfer from the first ball to the second at the moment of collision, as it transfers its momentum to the other ball: setting it off in a particular direction and at a particular speed. Though some of the "form" is transferred to the environment (e.g., transformed into sound or heat), some part of the information in the movement of the second ball is identical as in the movement of the first ball, meeting Collier's condition. Collier's position is best viewed as a metaphysically-oriented account of the nature of causation (considering how he posits it in respect to the other dominant views at the time), hence it might not be the most wieldy for analyzing causal relationships. It does offer an account of a "merely causal" processes, such as the billiard balls example from the list that opened this section, in terms of information flow.

Nevertheless, as I will distinguish in chapter 4, his reliance on the notion of *information content* (Kolmogorov complexity), rather than *transmission*, allows to still distinguish the case of billiard balls from the properly communicative processes. For now it is enough to note that communication—and, by extension, information transmission—requires the presence of a sender and a receiver. While in the ball case we have two distinguishable entities, they do not meet the criteria of activity identified by Levy. Further, the information in question is not identified in terms of the relationship between the entities, but rather a product of a description by an external observer. The latter point falls out from the discussions in chapter 4, so I bracket it for now. The former point is what requires some specification in naturalistic terms.

The notions of "action" and "agency" have long been a subject of discussion in philosophy of biology[21]. While standard accounts of action within the model of efficient causation usually led to the introduction of some special causal powers,

---

confusion, I will use the notion "NPI-content" to refer to information content in the sense of NPI, and "information content" otherwise.

[21] In what follows I decided not to include considerations from Philosophy of Action, as the field focuses on the high-level, intentional actions characteristic for humans and perhaps a few other species of animals. The definitions of action that have been advanced by such notable philosophers as e.g., Elisabeth Anscombe (Anscombe 2000) or Donald Davidson (Davidson 1963)

more recent attempts at a naturalistic description of "agent causation" (e.g., Bechtel and Bollhagen 2021; Potter and K. J. Mitchell 2022; Skewes and Hooker 2009; see also Bickhard 2000; Bickhard 2009) seem to converge on the special energetic properties of biological materials (Brancazio 2023; Harrison 2023; Harrison, Rorot, and Laukaityte 2022; McGivern 2020). What this emerging view highlights is that biological systems—organisms and cells in particular—are characterized by endogenous activity. In this context, actions are distinguished (according to the definition of Skewes and Hooker 2009, p. 292; this is motivated by the perspective of Rosen 2012) as activities "carried out as part of an orchestrated autonomous organism response to a situation that is framed by anticipative potentiation", where the concept of "anticipative potentiation" is intended to mark the goal-oriented character of action and its disconnection from the immediate, energetic setup of the biological system in which it is embedded (i.e., detachment in the sense of subsection 3.5.1). Crucially, Skewes and Hooker view action as a continuous modulation of an extended process, rather than a singular event.

At a cellular scale William Bechtel and Andrew Bollhagen (2021) propose to apply this conceptual framework by focusing on how Gibbs free energy in the form of ATP molecules is transformed into activities, such as molecular motor movement. Their account focuses on how ATP binds to myosin and then undergoes hydrolysis, producing chemical energy resulting in physical movement at the ATP binding site, which then travels along the myosin, changing its shape and resultingly—its action on actin. This is a relatively narrow example, but it highlights that despite its high-level, speculative framing and organismal focus, the model proposed by Skewes and Hooker can inform explanatory work within the new mechanist paradigm at the cellular scale.

Such an energetic view of action is closely related to the notion of "detachment" introduced above, as the endogenous activity or closure leads to some degree of detachment from the environment. Hence, the entities involved in communication require a particular organization, characterized in terms of their energetic properties, which enables them to *act*, whether it is through sending a signal or responding to one, in a way that is a physical causal process (as I will further specify in chapter 5), but a detached, non-deterministic (at least in the epistemological sense) one.

This is in fact captured by some of the definitions of communication. For instance, Jack Bradbury and Sandra Vehrencamp (1998, p. 356) in their definition claim that a signal "provides information used by another animal (the receiver) to select an action beneficial to both parties", which, as analysed by Artiga (2021, p. 8583), indicates that "there must be a range of acts or structures (at least two), such that receivers select their behavior." Effectively, the receiver (but also, the sender, as Artiga indicates that the action of sending the signal needs to be *a* possibility among a larger set—this directly follows from Shannon's definition of information, as I will show in the next chapter) must be detached to some degree.

---

unfortunately are not suited to explain biological activity, where the ascriptions of intentions is not a given.

If there is no choice of action, and effectively no detachment, the process would not qualify as communication under these terms.

As the organization required for detachment is characteristic of living systems, it allows for delineating the cases such as the billiard ball example from the genuine cases of communication in biological systems, precisely because billiard balls are inert, and their interaction is characterized by highly limited detachment. But this distinction is one of a degree, as clearly follows from the definition of detachment (as a measure of the relative independence), and as such it is likely to lead to some boundary cases, where intuitions break and the question if the process has enough detachment or not are difficult to address. This is one of the reasons why the account proposed here is not a full *definition* of communication, as in such cases the researchers' preference for communicative or causal language will likely differ and eventually settle on one of the descriptions, guided by both theoretical and theoretically external factors.

3.5.3 *Communication and natural signs*

Consider now the final two cases, of the digestive system (Case 4) and predator-prey tracking (Case 5), from the list at the beginning of section 3.5. For convenience, I will quote them in full:

- Case 4: The digestive system breaks down carbohydrates into monosaccharides, such as glucose, which get released into the bloodstream and stimulate the pancreas to secrete insulin. We have a transfer of a mark here, in the form of glucose, we have separate entities—the digestive system and the pancreas, we have a structure—the bloodstream, the specific receptors that pick out concentration of sugars. But it's not communication!
- Case 5: A predator picks up the scent of a prey. We have the transfer of a mark, the pheromones left by the prey animal, e.g., on their feces, we have distinct entities involved and a structure which evolved for this particular function. Yet again, it's not communication!

These cases, as highlighted in the description, do meet the criteria set out in the proposed definition-schema for communication. In both cases, we have distinct or distinguishable entities involved, a mark that is transferred between them and a structure which supports that transfer. Further, in both cases, it is the function of the putative receiver to pick out the mark. Yet, neither of those cases can be appropriately described as communication according to the standard use of the term. Instead, these are two instances of natural signs, or cues.

While prevalent, this distinction is not universally accepted. For instance, the *Encyclopedia of Evolutionary Psychological Science* (Freeberg et al. 2021) defines "communication" as including both the active signals and passive cues. But even in that context, the distinction of signals and cues is quite strict: these authors characterize signals as "displays or acts produced by senders that influence the behavior of receivers" because the acts and the responses have co-evolved (Freeberg et al. 2021, p. 1209), and cues as "phenotypic variation in a sender that causes a reaction in a receiver that benefits the receiver even though the phenotypic variation in the sender has not evolved to cause that reaction" (Freeberg et al. 2021, p. 1206).

How can we consider this distinction within the framework proposed here?

Let's begin by considering a broader category. The concept of a "natural sign" refers to events that are reliably correlated with each other so that in virtue of perceiving one of them, we can infer the occurrence of the other. The "reliable correlation" criterion works best when there's a causal relationship between the events: for instance, a column of smoke raising above a forest is a natural sign of fire, or fever is a natural sign of the flu. But causality need not be involved, as Ruth Millikan highlights in her discussion of natural signs (2012, p. 497): the North Star is a sign of the geographical north, or when "[t]raveling north from R[oute] 89 on Wormwood Hill Road, the pond on the right is a sign that our house is coming up next", though no causal relationships are involved in both cases.

This relationship is sometimes captured through terms such as "natural information", which highlights the fact that the correlational relationship between the events can be formalized in terms of mutual information (see chapter 4), or alternatively—as "natural meaning", to use a term introduced by Paul Grice (1957), which focuses on the semantic relationship between the events: what the natural sign is about (I explore the semantic side in greater detail in chapter 5).

While these causal or correlational relationships are facts about the world, most accounts of natural signs assume that their semiotic character is grounded in activities of some agents which interact with them. Millikan captures this by saying that natural signs result "from endurances and repetitions of patterns along paths that cris-cross and interweave with the paths of the animals for whom these signs are affordances" (Millikan 2012, p. 505; for "affordances," see Gibson 1986). This requirement of a receiver is what they have in common with communication: there is indeed a mark, and it is a function of the receiver to pick that mark—which is possible only if an appropriate structure is present.

Resultingly, what differentiates this case from actual biological communication is the absence of an *active sender*. Consider standard examples of natural signs. Me having a fever is not an attempt to signal to the doctor that I have the flu, similarly the prey does not leave its scent to signal its location to the predator, even if both the doctor and the predator are capable of identifying those marks and using them as signs of, respectively, the flu or the presence of prey. Some examples are more complex, however. For instance, in the example 4 from the list of boundary cases, the digestive system does actively release glucose into the bloodstream, and it is the function of the receptors in the pancreas to react to the sugar concentrations in the blood and release insulin appropriately. It is not, however, the function of the digestive system to signal the blood sugar level to the pancreas. If we consider a standard description of causes of (some variants) of type 1 diabetes (see the World Health Organization overview, WHO,), where the pancreas do not produce sufficient insulin, it is not treated as a dysfunction of the digestive system, but rather of the pancreas. Hence, we do not take the digestive system to perform the function of signalling blood sugar levels, but rather take the putative recipient—the pancreas—to perform the function of picking out the concentration as the sign for sugar levels.

Ascriptions of function to biological systems are notably problematic (I review this topic in section 5.1), though widespread in science. Indeed, "function" offers a relatively straightforward way to characterize the distinction of cues and signals: by stipulating the additional condition that the communication system requires both the sender and the receiver to have the function of—respectively—sending out and receiving signals. This condition is in fact quite widely accepted in the work on animal communication (e.g., Maynard Smith and D. Harper 2003). Nevertheless, defining the sender through their communicative function risks circularity, since the definition-schema proposed here requires the sender and the receiver. Hence, I will explore an alternative view, which does without the notion of "function."

This alternative is suggested by the definition of Freeberg and colleagues quoted above, namely that signals are "displays or acts" and cues are properties, or more precisely—"phenotypic variation." This suggests that the notion of "signals"—and "communication" more broadly—introduces some constraints onto what marks (in the sense of the definition-schema) can be appropriately considered. Broadly speaking, it is the distinction between active and passive characteristics of the putative senders (such a characterization of this difference has also been suggested by Butlin (2020), though his account focuses on the activity of the receiver). Importantly, this distinction has already been introduced above in my discussion of the billiard ball example (subsection 3.5.2).

The framework of organizational closure, discussed in subsection 3.5.2 and 1.2, offers some central criteria that can be used to distinguish activities from passive properties—and effectively, signals from cues. Consider the difference between the glucose example above (example 4) and the already discussed case of potassium signalling in biofilms (case 3), as both are closely related to metabolism. The digestive system releases glucose to bloodstream to provide and distribute required nutrients to the organism. This is an activity, as it is endogenously-driven and results from "orchestrated autonomous response." Nevertheless, the relevant "framing of anticipative potentiation" (Skewes and Hooker 2009) does not include the pancreas producing insulin—the digestive system is not directly sensitive to the amount of insulin.

On the other hand, in the case of potassium signalling, the release of potassium similarly is the result of an activity of the cell, and the presence of active propagation highlights that there is some "anticipative potentiation", especially since the ultimate goal of the signal is to reduce the growth at periphery so that the nutrients diffuse towards the center—to which the bacteria are very much sensitive. Effectively, then, what distinguishes signals from cues is that signals are those activities of the biological system which contain a response from some other biological systems in their "anticipative potentiation", i.e., they're sensitive to that response in some way. Marks considered in the definition-schema need to be outputs of those activities, and the structure of the communicative processes needs to enable this sensitivity to response.

The "sensitivity to response" is an important category that requires some further consideration. As I have mentioned before, in some cases, biologists wish to consider processes such as memory in terms of communication. If we have a spatially continuous system where the sender/receiver distinction is posited only in the temporal dimension, the bidirectionality of structure suggested above can be problematic, as the sender in the past cannot receive response from the receiver in the future (according to the contemporary best physics). This is unique to memory-like cases, as in all other cases we can consider the spatial distinction of the sender/receiver systems—meaning that the sender can exist to receive the response, whereas in the memory case, the reception of the signal by definition "destroys" the sender.

One way to account for that would be to consider the sensitivity in terms of functions (which, according to the dominant teleosemantic account—see chapter 5—are primarily attributed to the whole organism, and only derivatively to its parts), but the decision to focus on activity was motivated by the risks of circularity associated with applications of function.

Instead, I propose to consider the sender/receiver distinction as non-dichotomous, except for relative to a particular instance of communication. Indeed, the sensitivity to a response can make each sender *also* a receiver in some (other) communication process. Virtually all examples from the corpus consider communication between systems that can—according to the circumstances—act as both senders and receivers. Hence, the sensitivity to a response is a broader property, not indexed to any particular communicative instance. It is enough for the sender to anticipate *some* response (i.e., the response can be relatively distant in the causal chain; and such a response can itself not in fact be an instance of communicative interaction). Resultingly, in the memory-like cases we can consider the sender/receiver as a single system except when analysing a particular, temporal-only instance of transmission.

What follows from this view of sensitivity is that the signal/cue distinction only makes sense relative to a particular sender/receiver system, but this is a welcome result. Consider the eavesdropping in the example of *Photinus macdermotti* fireflies (based on the description in Stegmann 2009). The males of that species emit light flashes, to which females respond with their own pattern of flashes if they are willing to mate, allowing the male to locate them. However, there are species of fireflies which prey on other fireflies: one example is the females of the species *Photuris versicolor* which prey on males of several *Photinus* species, including *P. macdermotti*, by mimicking the flashes of females of those species and luring the males.

What happens in that situation is that a *P. macdermotti* male sends a signal to communicate—and eventually mate—with a *P. macderotti* female, but if a *P.*

versicolor female is present she may take the male's initial signal as a cue that there is potential prey around. In this case, a communicative situation is being hijacked and treated as a natural sign (relative to a different sender/receiver system). The action of the male—the light flashes he emits—are a signal to the female of its own species, but the presence of those flashes can be at the same time a cue from the perspective of the predator. The anticipative character of the adopted concept of action allows for this distinction.

3.6 SUMMARY

In this chapter, I have analysed the corpus usage of the notion of COMMUNICATION as identified by semantic search, using definitions from the literature as queries, and following the methodology laid out in chapter 2. The goal was to provide an account of how the concept is used across the various contexts it appears in. I have explicitly not attempted to formulate a definition of communication, but instead characterize the criteria that guide its application in scientific literature. Hence, the resulting account remains quite broad. Depending on a variety of theoretically grounded decisions, it can be further specified and reformulated into a definition. Nevertheless, the criteria are sufficiently detailed to apply to a variety of standard problems in the discussion of BIOLOGICAL COMMUNICATION. In this chapter, using several concrete cases, I have focused on the distinction between "merely causal" and genuinely communicative processes (subsection 3.5.1 and 3.5.2), as well signals and cues (subsection 3.5.3), which have allowed to further develop the details of the proposed view, as well as highlight its potential to solve outstanding problems in philosophy of biology.

According to the analysis, a biological process can be described in terms of communication if several criteria are met. First, the structure of the process involves:

(1) transfer of a *certain mark* K,
(2) activity of *distinguishable* entities, X and Y, and
(3) the presence of a *structure* S enabling such transfer or exchange.

The components must meet additional criteria:

1. The mark, K, is a physical cause or a sequence of causes, characterized by material discontinuity at least at the point of reception (i.e., when reaching the entity Y; as characterized in subsection 3.5.1).
2. The distinguishable entities, X and Y, actively participate in the process, which requires them to have some degree of detachment or organizational closure (as characterized in subsection 3.5.1 and 3.5.2):
    a. X, the sender, actively produces the mark;
    b. Y, the receiver, actively produces some response in result.
3. The structure, S, involves:
    a. the anticipative potentiation (understood in terms of sensitivity to response) of the sender, X (as discussed in subsection 3.5.1);
    b. energetic asymmetry between the signal and the resulting response, which specifies a direction (relative to a particular communication instance; see the discussion in subsection 3.3.1 and 3.5.2).[22]

---

[22] Additional specifications regarding the structure will follow from the analysis of semantics of communication chapter 5.

4. An individual instance of communication is embedded within a larger causal process, and as such it can be characterized as performing some function in that process (see chapter 5).

As highlighted throughout this chapter, this characterization of communication is shared across all the sc of biological communication at which communication is invoked in the corpus. Effectively, then, communication can be characterized as a scale-free process, as discussed in the Introduction. Further, the addition of the criterion 4 means that we can identify a strategy for scaling up communicative processes: communication at some scale can be a part of a larger process, which is communicative in virtue of the smaller-scale communication it involves, e.g., due to the detachment of components and sensitivity to response of the structure of that smaller-scale process. This is quite natural when we consider some familiar cases, for instance linguistic communication, that according to the current best model involves a plethora of communicative instances—nervous signalling in the brain, the neural control of motor actions—in virtue of which such linguistic communication is possible. However, this view has also some broader implications that I will explore in chapter 6. Before that, there are some outstanding questions in philosophical discussions of biology that I will explore from the perspective of the proposed model of BIOLOGICAL COMMUNICATION.

4

JUXTAPOSING BIOLOGICAL COMMUNICATION WITH FORMAL ACCOUNTS

While the emerging picture of BIOLOGICAL COMMUNICATION captures the scientific practice related to describing various processes across the scales of biological organization in communicative terms, as I hope to have shown by now, it lacks an important component. There is a pervasive connection in the corpus between COMMUNICATION and INFORMATION, despite some theoreticians' arguments that information talk offers no benefit in explaining communication (primarily Scott-Phillips 2008). The references to information, on the other hand, are introduced almost exclusively within the mathematical theory of communication put forward by Claude Shannon (Shannon and Weaver 1964). Effectively, information offers the dominant (and perhaps, the only) way to formalize and mathematically model communicative interactions. This formal background frequently appears only declaratively, as the real use of the term departs from the Shannonian framework. Nevertheless, it is necessary to consider it in detail to shed light on some assumptions that emerge from information theory and shape the understanding of communication across biology and cognitive sciences.

Hence, in the current chapter, I discuss this framework in greater detail, together with three formalisms that have been constructed on the basis of Shannon notion of information, and show how information theory connects to the biological use of "communication" and "information." As readers familiar with the formalism will have noticed by now, the model put forward in the previous chapter is compatible with Shannon's account. Contrasting it directly with the existing formalisms highlights several points of contention for biologists' reliance on information theory. I will explore this topic in detail below. While the primary goal here is uncovering the interdependences of COMMUNICATION and INFORMATION, there are two further epistemic benefits that the discussion will help achieve.

First, there's a large body of work in philosophy of biology that explores how "information talk", as this conceptual perspective has come to be known (e.g., Griffiths 2001, p. 394), shapes research in life and mind sciences since the 1950s (an area that I have already discussed in chapter 1). I will go back to this work in order to connect it to the corpus data, to explore how well it tracks larger-scale patterns of use of information theory, the accompanying assumptions, etc. This intersects with several important discussions. First, the problem of non-ergodicity (and non-stationarity) of living systems, which goes against important formal restrictions of Shannon information theory. Second, the problem of biological individuality which impacts the way we can distinguish "senders" and "receivers" in communication processes (a problem I have mentioned in the previous chapter).

Second, as I will argue, biological communication escapes some central assumptions of information theory, and the empirical data suggest a complex dynamic of how conceptual references to "communication" intersect with more formal treatments. This foregrounds the issues related to abstraction and idealization in science, a central research area in contemporary philosophy of science. Indeed, some previous theorists have explored this and argued for an anti-realist interpretation of information talk (e.g., Griffiths 2001; Levy 2011). In the thesis, I advance broadly realist interpretation of BIOLOGICAL COMMUNICATION, as discussed in chapter 1 and chapter 6, hence I will explore how these idealizations can be accounted for at some length.

These explorations allow us to introduce crucial nuance into the way BIOLOGICAL COMMUNICATION is treated in the literature and what epistemic roles this notion plays. In what follows, I will begin in section 4.1 by recapitulating the basic tenets of information theory and give a broad overview of the (biologically relevant) developments of the framework that followed Shannon's standard treatment, to ensure common ground. There are important differences in how information is conceptualized that follow from these formal treatments—differences often overlooked in the use of the term.[1] This section helps differentiate the different conceptualizations, as I will track them later on. Then, in section 4.2, I will turn to the exploration of conceptual implications of information theory for "communication talk" in the corpus. I will approach this question through the two narrower issues: the impact of information-theoretic assumptions and the role of the information theory formalism as an idealization. I conclude in section 4.3 by offering a systematic proposal of the role that formal information theory may play in biological discussions of communication.

---

1. While I'll argue that these differences are quite deep and heavily impact the conceptual practice of biology and cognitive science, all frameworks discussed in this chapter are concerned with so-called *syntactic* information, i.e., broadly correlational sense of the term, without presupposing any aboutness that characterizes *semantic* information (which is discussed in chapter 5). In that sense, all these frameworks closely follow Shannon's lead, who—as I'll discuss in subsection 4.1.1 below—cemented this agnosticism with regard to meaning for formal treatment of information.

4.1 AN OVERVIEW OF (BIOLOGICALLY RELEVANT) INFORMATION THEORY

In this section, I provide a brief introduction to information theory and highlight some of the directions it took since its inception. This is not a comprehensive review, and is not intended as such. Instead, I selected the formalisms that are most impactful for the life and mind sciences—they are referenced and used either in biological literature or meta-theoretic analysis of "information talk" that I have reviewed in chapter 1. As such, these conceptions will inform my analysis of the empirical data in the second part of the chapter. I omit some conceptions—for instance, James J. Gibson's notion of ecological information (Gibson 1950, 1966), Luciano Floridi's definition of information as "data + meaning" (e.g., Floridi 2010, 2011), or Fred Dretske's account (Dretske 1981, the two latter concepts apply to "semantic" information, and I come back to them in chapter 5)—as they do not figure in the debates about *biological* information. For instance, despite the relevance of Gibson's ecological psychology for cognitive science, his notion of information is rarely picked up in more recent discussions.

4.1.1 *Mathematical theory of communication*

The dominant formalization of the communication process has been offered by Claude Shannon (1948) and then developed in collaboration between Shannon and Warren Weaver. The two mathematicians worked on telephonic communication as the central case of information transfer. Their conceptualization of the problem was greatly influenced by Shannon's role as a cryptanalyst during the Second World War (see Gleick 2011, ch. 7) and the earlier work on telegraphic systems by Harry Nyquist and Ralph Hartley in the 1920s, his older Bell Laboratories colleagues (Shannon and Weaver 1964; see Gleick 2011, ch. 6). Shannon coined the modern meaning of the term "information". Following Nyquist and Hartley, who sought to "eliminate the psychological factors involved" from the understanding of communication, and turn it into a purely engineering task, Shannon made the central step of tying this non-psychological concept of "information" to uncertainty, surprise, and—eventually—entropy.

In this probabilistic framing, the problem of communication is that of "reproducing at one point either exactly or approximately a message selected at another point" (Shannon and Weaver 1964, p. 31). As such it can be modelled by the (now standard) diagram presented in Figure 4.1. The model assumes 5 central elements (quoting from Shannon and Weaver 1964, pp. 33-34): (i) the *information source*, "which produces a message or a sequence of messages"; (ii) the *transmitter*, which "operates on the message […] to produce a *signal*" which is then transmitted over (iii) a *channel*, "merely a medium"—which offers the physical means of conveying the signal to (iv) the *receiver*, the inverse of the transmitter, which reconstructs the message so that it can reach its (v) destination, "the person (or thing) for whom the message is intended".

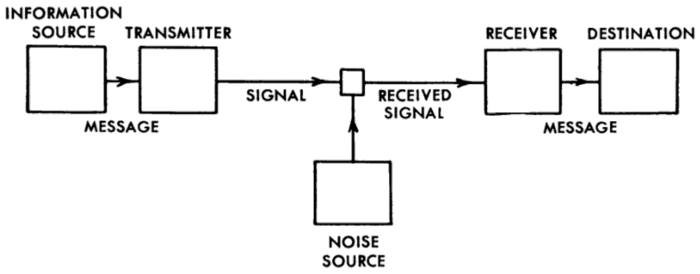


Fig. 1. — Schematic diagram of a general communication system.

Figure 4.1: Shannon's model of communication, as presented in Shannon and Weaver (1964).

As mentioned, this model is motivated by the technological means of communication of the time: when speaking over a wired telephone I, the information source, am speaking into the microphone, the encoder, which transforms sound waves into electrical pattern and transmits it through a (complex system of) wire(s), the channel, to my interlocutor's phone, the receiver, which transforms the electrical pattern back into sound waves so that my friend hears the words I'm saying.

But this example abstracts away a lot of important intermediate steps, so for a fuller picture consider a simple case of sending a message consisting of a string of 0s and 1s of a fixed length N over a wire. What the information source actually does is they pick a particular message out of the set of all possible messages: a particular binary string out of the 2^N possibilities. This message is translated into a signal, in our case, a sequence of current or voltage fluctuations in the wire, which are picked out by the receiver. The problem for the destination is then to match the electrical pulses to 0s and 1s appropriately: especially in the "normal" circumstances where the transmission is imperfect in unknown ways, to an unknown degree.

Since any message in real-world circumstance is essentially sequential, hence, as Shannon suggested, it can be modelled as "the outcome of a process that generated events with discrete probabilities" (Gleick 2011, p. 228, with the caveat that in some cases a continuous treatment is necessary). In that context information can be measured as the amount of "of how much 'choice' is involved in the selection of the event [in the stochastic process] or of how uncertain we are of the outcome" (Shannon and Weaver 1964, p. 49). While the probabilistic structure of the message may vary—depending on what is transmitted, e.g., languages have different distributions of probabilities of letter sequences—Shannon was able to work out a single measure of the uncertainty involved, H, that became known as "Shannon entropy":[2]

---

[2] Shannon was inspired by the thermodynamic conceptualization of entropy in his work, and the definition parallels that of thermodynamic entropy, hence the name. However, this parallel introduces a broader and still open problem of the physical basis of information. The

$$H = -K \sum_{i=1}^{n} p_i \log p_i$$

where K is a positive constant, and p_i is the probability of the i-th message from an n element set.

Consider again the sequence of 0s and 1s of a fixed length. Assume now that the message transmits a result of N fair coin tosses, so that 0 corresponds to heads, and 1 to tails. In that case, any position in the sequence can contain 0 or 1 with equal probability, $\frac{1}{2}$. The set of all possible messages grows quite quickly with the length of the message, $n = 2^N$, and the probability of each message is equal: for any i, $p_i = \frac{1}{2^N}$. If we have a message of the length 1 (ignoring the K, which selects the unit of measure), the total information of any message in our set will be:

$$H = -\frac{1}{2} \log \frac{1}{2} \times 2 = -\log \frac{1}{2} = 0.301$$

A logarithm of base 10 is rarely used for calculation of the amount of information—the most popular is the base of 2, which defines the ubiquitous measure of *bits*, very useful for quantifying the amount of information carried by a single binary digit in a sequence of the sorts considered here:

$$H = -\log_2 \frac{1}{2} = 1$$

Essentially, the exact value is meaningless. What matters is the relationship between the amount of information and probability: the more certain a particular message is, the less information it carries—as it resolves less uncertainty. The 0s and 1s in our string carry the most information—exactly 1 bit each—as long as they're equiprobable.

The probabilistic framing of information was developed concurrently with the work on the foundations of computer science. Crucially, the early information theory coincided with Alan Turing's mathematical breakthroughs, whose cryptanalytical period during World War II led him to similar conclusions as Shannon's, though he developed them in notably different ways (see Good 1979). Almost equally importantly, Shannon's theory was quickly received by the emerging cybernetic community in the United States of America and in the United Kingdom, even though Shannon's relationship to Norbert Wiener—a famously difficult person (see Gleick 2011; Pickering 2010) who nonetheless shaped the cybernetic movement (Wiener [1948] 2019)—was strained. Regardless of their personal tensions—although the exchanges James Gleick catalogues in his book

---

transmission of information necessarily requires some form of expenditure of energy (in computation, Landauer's principle (1961) defines the lower bound of this value as proportional to the temperature of the system), but the details of this relationship are complex. This relationship is explored in some detail by Terrence Deacon (2007; 2008) and by John Collier (2011). The material basis of communication will turn out crucial in the following chapter 5, though I will not be able to treat this topic at sufficient depth.

(2011, ch. 8) serve as a very good introduction to old-school academic passive aggression—cybernetics quickly became intertwined with information theory and led to a series of models critical to this day. These include: (1) the model of a neuron as an electrical switching mechanism offered by Warren McCulloch and Walter Pitts (McCulloch and Pitts 1943), (2) John von Neumann's work on computation and his introduction of game theory and cellular automata model (Von Neumann and Morgenstern [1944] 2007; Von Neumann 1966), and (3) eventually, after the structure of the DNA was uncovered through Rosalind Franklin's pioneering work in X-ray crystallography and the analysis of her results published by John Watson and Francis Crick (Watson and Crick 1953), the description of the function of the genetic code (Gleick 2011, ch. 10). It was also through cybernetics (see Gleick 2011, ch. 8) that information theory was connected with cognitive psychology in the work of Donald Broadbent, George Miller, and Herbert Simon. The concept of "information processing" began to replace the behaviourist language of stimulus and response, and shaped the then-nascent discipline of cognitive science. These developments accompanied a much broader spread of information-theoretic language even though Shannon became quickly wary of the oft-superficial interpretations his framework has been receiving across a growing number of research fields (see Shannon [1956] 2021).

The connection of information theory to biology (in genetics and neuroscience) as well as to cognitive science (through computation) continues to this day and has generated a wide interest among philosophers of science as well. I have already reviewed some of these philosophical debates regarding information in biology and cognition (in section 1.3), and will go back to them below (in section 4.2) to analyse them in greater detail. But information theory has not been "completed", whatever this could mean, by Shannon, even if sometimes it is treated as a monolith of sorts, and has seen major developments over the course of the almost 80 years since Shannon's paper. These developments influence the various interdisciplinary applications of information theory and are critical for understanding how communication can be formally described.

4.1.2 *Assumptions of information theory*

For the formalism to be mathematically tractable, Shannon needed several assumptions that are quite standard in the study of stochastic processes (even if they misrepresent some target systems in critical ways, as we will see below). I will review two sets of assumptions briefly here to introduce key notions, as they will be crucial for understanding the role and limitations of information theory in life and cognitive sciences (see section 4.2).

Sender/receiver model

As noted above, Shannon uses a 5-element model to introduce his formulation of the problem of communication. This model is structurally quite similar to the one that I have arrived at in the previous chapter (further highlighting the importance of his formalism in the literature). It is, however, more fixed or determinate: in particular, the strong delineation of the information source and the transmitter (and, analogously, the receiver and destination) as well as the conceptualization of the "channel", are strongly embedded in the specific engineering context Shannon was working on: ensuring successful information transfer over a noisy channel (in telephony, but this easily translates to related domains: television, telegraphy, or radio transmission that became dominant later on). While some aspects of the description can often be abstracted away (e.g., the channel can be described in terms of noise and capacity, instead of specifying a particular physical medium it is made of), the formulation of this 5-element sender/receiver model does introduce requirements that are demanding for some natural systems.

Most importantly, as I've shown before, the sender/receiver model connects closely to the formulation of the problem of communication in Shannon's framework. On the risk of being redundant, for Shannon, information is the reduction of the uncertainty at the receiver, regarding the message sent by the source, based on the signal that reaches the receiver. *Information* in this model can be measured only from the perspective of that sender/receiver system. The formalism, however, turns out (in the course of further developments) to be more powerful than that. In some cases, the use of the formal notion of information—in terms of "content" or "flow", see below—does not require the sender/receiver model. Nevertheless, the use of the information *transmission* formalism to model biological systems requires biting the bullet on the description of that system in terms of the sender/receiver schema. This poses some important problems for understanding the nature of scientific knowledge created with use of this formalism (see subsection 4.2.1).

Markov condition and the ergodicity of the information source

The probabilistic framing of communication means that the information source can be modelled as a stochastic process specified by a probability distribution. Shannon assumes—and the mathematics of his theory of information transmission crucially depend on this assumption—that the process in question can be modelled as a discrete Markov process, and even more restrictively, Shannon focuses on the class of ergodic Markov processes.

A Markov process, described by a Russian mathematician Andrey Markov, is a stochastic process that generates a sequence of discrete events with the property—known as the Markov property—that the probability of each state in the sequence depends only on the state attained in a previous step. Such "memoryless" processes include, notably, Brownian motion, a movement of a particle which at each step moves in a random direction, so that its position at step $t - 1$ depends only on its previous position at time t.

Ergodic processes, now well-studied beyond the domain of Markovian processes, can be characterized in several ways. In the context of stochastic processes, ergodicity is the property where the process time average (the average of states over a sufficiently long sample of a single sequence) is equal to that process ensemble average (the average of states at any single time step across multiple "runs" of the process). Alternatively, if we take the ergodic stochastic process to be a description of a dynamical system in a particular state space, that system will eventually (over infinite time) visit all parts of the state space.

Ergodic processes have important characteristics that make them extremely useful: for instance, if we measure the states of the system over a particular time window, the descriptive statistics of that measurement (in the limit, as the time window expands) will be the true descriptive statistics of the system.[3] Brownian motion, noted above, is indeed an ergodic Markov process: if we observe long enough trajectory of the movement, we can deduce the true statistics of the individual shifts of position at each time step, so the probability distribution of direction and speed.[4]

For Shannon, this assumption made sense, as the ergodic Markov processes are special cases of particular interest for probability theory where they have garnered significant interest and have a number of powerful properties identified that he could make use of. Further, in the engineering context he was interested in, these assumptions most often hold up at least approximately. But these special cases have limited applicability beyond that engineering context. For instance, the assumption of ergodicity does not apply even to the case of language that Shannon was interested in (see Dębowski 2021), and even more so in a strictly biological context (see e.g., Kauffman 2019; Medaglia et al. 2011; Weron et al. 2017). This leads to several issues in the case of biological information transmission, as we will see below.

On a conceptual level, the assumptions of sender/receiver model and the ergodicity of information source are closely connected: both follow from Shannon's focus on the efficient coding in transmission over a noisy channel. In such cases, the formulation of the problem achieved by his model, and the restriction to ergodic Markov processes significantly limits the computational difficulties involved. More recent developments in information theory make do without those assumptions, as they shift the interest away from information transmission, especially in the context of computation (as I will discuss below). But this step restricts the ability of information theory to function as an account of communication, as I will argue in subsection 4.2.1.

---

[3] Note that this property relates closely to the fact that all ergodic processes are stationary, i.e., that their statistical properties (the underlying probability distribution) remain constant over time.

[4] To be precise, the position of the particle following Brownian motion is ergodic for bounded open sets in one- or two-dimensional real number space, $\mathbb{R}^1$ and $\mathbb{R}^2$, but non-stationary and non-ergodic in 3 or more dimensions, see Ugbebor (1980).

As mentioned above, while the references to information in the corpus frequently invoke Shannonian information theory, most often they do not rely on that formalism directly. Instead, they draw on the various developments of the theory, which significantly modify the conceptualization of key notions, including "information," in philosophically important ways. Hence, to highlight in what respect information theory can act as a formal account of biological communication, in the following subsections, subsection 4.1.3 and subsection 4.1.4, I attempt to provide a broad overview of the different formalisms that emerged from the basis of Shannon's original formulation of mathematical theory of communication. This allows me to capture the conceptual shifts in the field, which hitherto have not been explicitly discussed in philosophical literature. Eventually, this will provide further specification of the concept of BIOLOGICAL COMMUNICATION.

4.1.3 *Information content, information flow*

Shannon's interest in transmission and efficient coding significantly shaped the format of information theory: it is notable that while he offered it as a "mathematical theory of communication" it got picked up as "information theory". Indeed, Thomas Cover and Joy Thomas in their important textbook treatment (2006, pp. 1-4), note that "communication theory" makes up just a portion of information theory, along with physics, statistics, mathematics, probability theory, economics, and computer science.

In the context of the more contemporary use to information theory, information is understood through the notion of "mutual information".[5] Initially introduced by Shannon as the rate of transmission over a discrete, noisy channel (see Shannon and Weaver 1964, p. 67-68) the now-standard term of "mutual information" has been proposed by Robert Fano (see Kreer 1957). "Mutual information" is the reduction in uncertainty in one random variable due to another variable—a special case of relative entropy, which measures the difference between two probability mass functions (Cover and Thomas 2006, pp. 6-9). The standard definition for two random variables X and Y is:

$$I(X;Y) = H(X) - H(X|Y) = \sum_{x \in X, y \in Y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}$$

Note that this is a symmetric relationship. The conditional entropy $H(X|Y)$ is defined analogously to absolute entropy $H(X)$ with conditional probabilities $p(X|Y)$:

$$H(X|Y) = \sum_{x \in X, y \in Y} p(y,x) \log \frac{p(y,x)}{p(y)}$$

---

[5] This point has been raised by Terrence Deacon in conversation after his seminar at the Human Interactivity and Language Lab, University of Warsaw, in September 2024. For Deacon, the Shannon entropy remains important, but it serves to define a *signal*, rather than information. This distinction is also explicitly recognized and discussed by John Collier (2008).

where $x \in X$ and $y \in Y$ describes all possible values of the random variables considered. This function describes a relationship between the probability distributions of two random variables. As such, mutual information does not require specification in terms of transmission (although it can be specified in those terms, as per Shannon's original definition). Effectively, it is a measure of information content: of how much information one variable carries about another.[6]

This approach to information theory as theory of *content* has been further developed in the context of data compression problems. This perspective is closely related to the problem of efficient or optimal codes Shannon explored, but ties to the advances in formal computer science. In that context, several researchers—Ray Solomonoff, Andrey Kolmogorov, and Gregory Chaitin—working at the intersection of probability theory and computer science, independently developed a framework that became known as algorithmic information theory (AIT). The goal of AIT was to develop "an objective and absolute notion of information in an individual object" (Hutter 2007), solving precisely the "problem" of dependence of Shannon information on the transmission schema, as I highlighted above. Further, in this way they were also able to disentangle the information theoretic formal apparatus from its reliance on *types* and, resultingly, *averages*, rather than individual occurrences. Strictly speaking for Shannon, notions such as "entropy" or "information capacity" of a channel only made sense in the average, for a particular communication system, and not for individual messages. In the framework of AIT, "information content" of a particular object (a string, for simplicity) is traditionally defined in terms of Kolmogorov complexity, which is the length of the shortest computer program that generates the string (assuming some reference universal Turing machine which computes the program). It is the shortest possible (formal) description of the object. The more complex a given string is, the larger its information content. AIT proved critical in formalizing randomness and probability of individual objects, leading to important insights across different fields of mathematics. In theoretical computer science it shaped the study of computational complexity and, more practically, allowed the development of widely used data compression methods.

While AIT departs from Shannon's formalism, the notion of Kolmogorov complexity is in fact mathematically related to Shannon entropy, as Ming Li and Paul Vitányi (1993, p. 141) prove: for a particular string, its *expected* algorithmic complexity is equal to stochastic entropy up to an independent, additive constant. It is a "beautiful fact" as Li and Vitányi note, and it highlights the continuity between AIT and Shannon's formalism. In that context Shannon entropy can be regarded as a measure of "frequency regularity" (M. Li and Vitányi 1993,

---

[6] The notion of "information content" in natural language can be read either in terms of information stored in an individual signal, or semantically, as what the signal means. To avoid confusion, in this chapter I use "content" only in the former, more technical sense of algorithmic information theory, as explained in this section, and reserve the notion of "aboutness" for the other sense of the word. However, the notion of "content" is so embedded in semantic theories that it was impractical and confusing to remove it from the discussions in chapter 5, where the default sense of the word is semantic.

p. 142) for individual strings generated by an information source.[7] Hence, even when investigating the relationship to Shannon's formalism, the algorithmic information theory obviates the Shannon model of communication: as there is no transmission, it doesn't make sense to look for the sender and receiver, or designate a channel. Similarly, in most circumstances, the assumption of ergodicity is not really needed, since the AIT doesn't require conceptualizing objects in terms of stochastic processes. In turn, it inherits a host of assumptions and restrictions from the selected formalism for computation: most importantly, as Gregory Chaitin has proved, the Kolmogorov complexity of a string becomes non-computable very quickly (as the length of the string increases).

This area has been particularly impactful in neuroscience, where the computational lens AIT requires is commonly accepted. In this context, the theory is variously used on a methodological level, as a way of examining the capacities and limits of various models (their computational complexity, or more precisely—its approximation), or as a method of data analysis. The latter is especially the case for neuroimaging or EEG studies, where information complexity measures are used for measuring neural signal diversity (e.g., used for distinguishing states of awareness, Schartner et al. 2017; Muñoz et al. 2020; for a review, see Turkheimer et al. 2022).

The importance of this shift from information transmission to information content (understood either as mutual information or algorithic complexity) can't be understated.[8] Many notable applications of information theory to life and cognitive sciences, some of which I have reviewed in section 1.3, rely on algorithmic information theory (e.g., Pocheville (2018); some do so implicitly). If that is the case, evaluating them from a purely Shannonian perspective (as is sometimes done in philosophical literature, see the discussion in chapter 1) is simply incorrect.

Nevertheless, the transmission perspective has not been abandoned entirely—among others, it found application at the intersection with epistemology, initially

---

[7] When introducing Shannon information theory, Li and Vitányi (1993, p. 63) make the following terminological note: "In information-theoretic terminology it is customary to say that the messages are produced by a 'stochastic source' which 'emits symbols' a_i with given 'probabilities' p_i. With abuse of terminology and notions, henceforth we use '*probability*' for 'frequency'. (Under certain conditions on the stochastic nature of the source this transition can be rigorously justified.)" If we take "frequency" in its regular meaning (as they do not provide any detailed, formal definition of the notion), it should be clear from the definition of ergodicity in subsection 4.1.2 that "certain conditions" mentioned include the ergodicity of information source. Even though their monograph doesn't reference the notions of "stationarity" or "ergodicity", these are sufficiently standard assumptions in mathematical contexts (and not directly relevant for AIT which is the main topic of interest for Li and Vitányi), that this implicit reference could have been considered sufficient.

[8] I owe this point to conversations with Carina Curto and Fernando Rosas during two consecutive editions of the School of Ideas in Neuroscience in Warsaw, in 2023 and 2024. Their points made me sensitive to this shift, which in the literature often happens implicitly, and I draw the content/transmission distinction directly from Rosas' characterization of the history of information theory (personal communication).

through the work of Fred Dretske (that I will review in greater detail in the next chapter, see Dretske 1981) and then in the formalization of the concept of "information flow" offered by Jon Barwise and Jerry Seligman (1997), who sought to explain the role of the notion of information for our understanding of knowledge and other epistemic phenomena, continuing Dretske's project.

For Barwise and Seligman, the notion of "flow" is meant to replace "transmission". Theirs is a theory which is meant to capture both communication and natural signs (as discussed in subsection 3.5.3 and 4.2.4). Indeed, they note that any "physical system whose behavior is modeled in probabilistic terms can be regarded as an information system in which information flows according to the equations of the [information] theory" (Barwise and Seligman 1997, p. 14). In that context, they focus on "distributed systems", divided into parts (they give example of a torch, distinguishing the lightbulb, batteries, the switch, and the case).

Their model introduces several key developments: for instance, they conceptualize the notion of the channel not as a physical medium for signal transmission, but as the regularities in the system—the set of "local connections" between system components defined formally in terms of infomorphisms between classifications, i.e., "whole-part relationships between instances of a whole [...] and that of a part" (Barwise and Seligman 1997, p. 32). In this way, they're able to tie the Shannonian focus on types ("classifications" for Barwise and Seligman) and averages, to the ability of particulars to carry information. The notion of infomorphism allows them to capture relationships between classifications, in virtue of which they can carry information about one another (building on the concept of mutual information). This will turn out to be a useful formalism for further distinguishing communication and instances of natural signals. Hence, I will specify this account in greater detail, in particular, their definitions of "classification" and of an "infomorphism."

A *classification* $\mathbb{R}$ is a triplet $\langle R, \Sigma_{\mathbb{R}}, \vDash_{\mathbb{R}} \rangle$, where R is a set of objects to be classified, called tokens, $\Sigma_{\mathbb{R}}$ is a set of objects used to classify the tokens, called types, and $\vDash_{\mathbb{R}}$ is a binary function between R and $\Sigma_{\mathbb{R}}$ which assigns tokens to types (see Barwise and Seligman 1997, p. 28, notation adapted). An *infomorphism* is a pair of functions $\langle f^{\Sigma_{\mathbb{R}}}, f^{C} \rangle$, such that for two classifications $\mathbb{R} = \langle R, \Sigma_{\mathbb{R}}, \vDash_{\mathbb{R}} \rangle$ and $\mathbb{C} = \langle C, \Sigma_{\mathbb{C}}, \vDash_{\mathbb{C}} \rangle$, and for any type $\rho \in \Sigma_{\mathbb{R}}$ and token $c \in C$: $f^{C}(c) \vDash_{\mathbb{R}} \rho \equiv c \vDash_{\mathbb{C}} f^{\Sigma_{\mathbb{R}}}(\rho)$.

Informally, an infomorphism is a pair of functions that "translate" between the classification of the whole system, $\mathbb{C}$, and of its components, $\mathbb{R}$. To illustrate this, Barwise and Seligman consider the example of a torch, the whole system, and its bulb, the component. The tokens are time-slices of the torch and the bulb: in that context, an infomorphism can assign to an instance of a torch the instance of the bulb at the same time (the function f^{C}), and assign to each type of the component (e.g., in a simple, everyday classification: $\Sigma_{\mathbb{R}} = \{\mathsf{LIT}, \mathsf{UNLIT}\}$) the corresponding type in the classification of the whole system (e.g., a more complex characterization in terms of physical properties; the function $f^{\Sigma_{\mathbb{R}}}$)

This formal apparatus allows for analysing the system and the relationships between its parts to pick out relevant regularities. Effectively, Barwise and Seligman strip down Shannonian model of communication to just the channel: the information source and the destination is captured as the classifications that are relata of the infomorphism, but their role is sufficiently captured just in terms of this relationship. This is also different from the definition of mutual information discussed above, as the infomorphism preserves the structure of the classifications and effectively defines the structure of the information flow, rather than quantifying the probabilistic dependencies of the classifications. Effectively, there is no "message" or "signal" beyond the relationships captured in terms of infomorphisms.

Their theory of information flow—the groundwork for a domain-general science of information, as they suggest—builds on the four main principles (Barwise and Seligman 1997):

1. "Information flow results from regularities in a distributed system." (p. 8)
    - This notion of regularity, recurring in the principles, is not rigorously defined in their account. It derives from the standard meaning of regularities as repeated events in a fixed pattern. Some instances of regularity may be nomic, conventional or logical (Barwise and Seligman 1997, p. 9). Importantly, they distinguish the "genuine" regularities from "merely accidental", though again—they are not providing specific criteria.
2. "Information flow crucially involves both types and their particulars." (p. 27)
    - Types and particulars are defined here in terms analogous as in the definition of classification above.
3. "It is by virtue of regularities among connections that information about some components of a distributed system carries information about other components." (p. 35)
4. "The regularities of a given distributed system are relative to its analysis in terms of information channels." (p. 43)
    - Principles 3 and 4 are cashed out in terms of classifications and infomorphisms: in the case of a torch, the regularity is that once we flip the switch on, the lightbulb lights up, because of the wires connecting the switch and the bulb (and the battery, obviously). Hence, the switch carries information that the bulb is lit. But if we pick a different information channel (a family of infomorphisms), e.g., consider the connection between the switch and bulb for a torch with a dead battery, the switch will no longer carry the information about the state of the lightbulb (as regardless of the switch position, the bulb can't turn on with a dead battery).

Barwise and Seligman's theory is intended to capture epistemological phenomena and refine Dretske's (1981) use of information for theory of knowledge, ultimately tying Shannonian information to the "everyday" notion. As such they develop it to cover a range of concepts in philosophy of science: the notion of scientific theories or inference (in local logics); as well as problems in philosophy of language: speech acts or semantic vagueness. Their model introduces also a notion of "aboutness" of information which will be discussed in detail in the next chapter. But the tools they introduce, in particular their "principles of information flow," will also prove important in capturing several aspects of biological information (as we'll see below, in subsection 4.2.4 and in the next chapter, in subsection 5.2.2).

4.1.4 *Signalling games*

Another crucial direction that accounts of information *transmission* took, was through the intersection with game theory and the introduction of "signalling games" by David K. Lewis (1969). Standard Lewisian signalling games, called sender-receiver games, aim to capture situations where agents with asymmetric information need to coordinate their actions. There are two players, a sender and a receiver. The sender knows the state of the world, the receiver doesn't, but needs to take an action on behalf of the pair. The sender can send a signal, based on which the receiver needs to choose the action. Then the pair receives a payoff. In the simplest version, there is just one "correct" action associated with each state that results in a symmetric, positive payoff. In response to any other action in that state, they don't receive any reward (for a broader overview see O'Connor 2020; Skyrms 2010).

If we assume such simplest scenario with exactly two states, two possible signals and two possible actions, the setup is straightforward. There are possible one-to-one mappings between signals and actions, which act as game-theoretic equilibria if the agents are rewarded for successful coordination. However, even this simple setup has been used to study e.g., dishonest signalling. Further, the general schema can be easily complicated to model a range of real-world situations: linguistic behaviour, including the evolution of language, economic decision-making, and a variety of biological—evolutionary and ecological—phenomena.

Lewis himself needed the concept to account for the emergence of convention in language and its role in establishing the semantics or aboutness of communication (which is explored in greater depth in the next chapter). While Lewis relied on the everyday notion of "information", these varied applications led researchers across the domains to study extensively both the probabilistic and informational properties of signalling games. Important work in this regard has come from Brian Skyrms (2010), who integrated Lewis framework with Shannon's formalism (see Zubek, Korbak, and Rączaszek-Leonardi 2024), as well as Carl Bergstrom, Simon Huttegger and Kevin Zollman (e.g., Bergstrom and

Lachmann 2003; Falk et al. 2025; Huttegger 2007; Meacham, Perlmutter, and Bergstrom 2013; Zollman, Bergstrom, and Huttegger 2013), who have applied signalling games to a range of phenomena in epistemology, philosophy of language and of science, as well as social and political philosophy, extending the model to allow for exploring of evolution and learning in these scenarios (see the overview in Skyrms 2010).

As a model of communication in biological systems, signalling games focus rather on *what* is communicated and *why*, than *how* that is of central focus in this thesis. But as such, they offer important insights into the role of communication: Skyrms states that "[s]ignals inform action, and signaling networks co-ordinate action. Signaling is a key ingredient in the evolution of teamwork" (2010, p. 3), an observation that we'll see also in the empirical data explored in the next section. The success of this model hinges on our understanding of the particular situation as one of information *transmission*.

While the Barwise and Seligman's account of information flow underscored the importance of the channel, here it is not directly modelled (a part of the *how* that is ignored in this approach) and the emphasis is put on the senders and receivers, their goals, knowledge, and abilities. These notions are traditionally introduced in higher-level terms, but with the assumption that they can be naturalized easily—especially the "goals" are in most biological cases regarded through evolutionary lens (see Bohl et al. 2014; Hummert et al. 2014).

In more recent developments, as complex games tend to be difficult to describe numerically, the researchers tend to use agent-based modelling to examine the possible trajectories of various signalling games. One such area is the studies of language emergence and evolution (see Zubek, Korbak, and Rączaszek-Leonardi 2024), where various signalling game scenarios are used to study the constraints impacting the development of the properties of language (e.g., compositionality, see Korbak et al. 2019, 2021; Kuciński et al. 2021). In these studies it is the *how* of communication that becomes central, highlighting the flexibility of signalling games.

Another example of this flexibility is offered by the signal/boundary framework proposed by John Holland (2012), and intended as an overarching theory of complex systems. Holland highlights that what cuts across the vast diversity of complex adaptive systems is the presence of boundaries, which delineate the systems' components (this is similar to the concept of "distributed" systems that Barwise and Seligman focus on), and signals, which pierce through the boundaries, connecting the components (more precisely, in the presence of signals, Holland offers to think of the boundaries as semi-permeable, as biological membranes tend to be).

He attempts to cover a wide range of systems: his central examples are the intricate web of ecological relationships in a rainforest ecosystem, the interactions at a stock market, or the protein exchanges between cell compartments. Holland's goal is to offer a formal description that applies across these domains. He does so by relying on the classifier system model of computation (for "defining signal-processing programs", Holland 2012, p. 25), urn models of probability theory (drawn from chemistry, used as models of semi-permeable boundaries) and genetic algorithms for accounting for their temporal evolution. While his model is perhaps more suggestive than strict, it reiterates several themes that recur throughout the other frameworks discussed here—as well as in the empirical results from the corpus, as we will see below. It will turn out highly useful in synthesizing the results in chapter 6.

4.2 TRACKING INFORMATION IN BIOLOGICAL COMMUNICATION

As we saw in the Introduction (chapter 1), biological—and even more so cognitive science—literature is replete with references to information. However, the applicability of formal information theory has been in the past a matter of debate for philosophers (as reviewed in section 1.3): the majority of participants of that earlier debate opted for distinguishing "semantic information", focusing on what the information is about, from the "syntactic" or "statistical" information that's the focus of the formal frameworks introduced above. As I've already mentioned in discussing the function of communication in the previous chapter, there are in fact important references to meaning across the scales of communication—a topic I'll analyse in detail in the following chapter. Nevertheless, information talk in biology involves also explicit and implicit references to a host of concepts embedded in the framework of information theory. It draws heavily on technical terms from that area, even if those terms are treated quite loosely at times. As highlighted in my previous discussion of this literature, this has been in fact noted by philosophers as well. Indeed, those early rejections of formal information theory in the biological context are motivated by significant and persistent misconceptions about the Shannon framework in the literature, both on part of its adopters and opponents.

This point has been raised among others by Stephen Mann (2020).[9] Mann raises accurate criticisms regarding the past mentions of the Shannon information theory in philosophy of biology, as too selective, missing important commitments of the framework, namely the transmission model. The most radical example of those mistakes in early philosophical work comes from Griffiths' (2001) paper, where he defines information content as "the amount of order in that system, or the inverse of the [thermodynamic] entropy (disorder) that all closed physical systems accumulate over time" (p. 396). This is not part of the Shannon framework to which Griffiths refers in this paragraph. Shannon did not in fact define any notion of "information content". It can be read as an incomplete and misleading interpretation of the Negentropy Principle of Information, as discussed in subsection 3.5.2 [Schrödinger (1944); Brillouin (1962); I'm following the analysis in Collier (1999); who also points out this kind of misreading of Shannon entropy as thermodynamic entropy]. Other authors also

---

[9] Mann's own proposal is set in a relatively idiosyncratic interpretation of the Shannon-Weaver framework. For instance, he treats signals as "symbols that stand for symbols" (Mann 2020, p. 670) of the original message. Strictly speaking, in Shannon's framework the original message doesn't need to be symbolic. Mann further offers a notion of a meaning of that signal as "an instruction how to recover the original symbol string" (p. 671). Mann likely refers here to the Shannon's understanding of communication as the task of "reproducing at one point either exactly or approximately a message selected at another point" (Shannon and Weaver 1964, p. 31). But his interpretation essentially ignores all the physical work that goes into establishing a communication system and that system's pre-existing setup necessary to treat a signal as such an instruction (e.g., Gleick 2011; Deacon 2011). More precisely, his point only makes sense within a very narrow, ahistorical focus on individual communication instances. Nevertheless, while I do not agree with his positive conclusions, I do believe that he raises important critical points.

confuse the relevant relata of the definition of information—e.g., Sterelny (2000, p. 196) suggests that information theory "defines information as the covariation between a signal and a source", disregarding the receiver.

At least in part, the incomplete application of Shannon's framework to biological systems comes from the fact that those systems do not lend themselves easily to the description with Shannon's model. Consider the context of genetics, where—as we saw in the introduction—information talk is applied widely, and has been a topic of extensive investigation of philosophers. If we think of the DNA as carrying information to be used in development, there is no clear sender, no clear message that is encoded in the nucleotide chain. This is the case even if the interpretation of the developmental system as the information reader (as proposed by Sterelny 2000) is quite convincing. While some philosophers attempted to reformulate the genetic phenomena as transmission, a more obvious route goes through focusing on the notion of information content, i.e., the information in an individual signal (see above in subsection 4.1.3). This route however has been noticed in that debate only by Pocheville (2018), while the Kolmogorov complexity would be a good candidate concept for the debates surrounding the uniqueness of the genetic causal pathway, both for the proponents of that view and the opponents advancing the causal parity thesis (see the discussion in chapter 1). Positing genes as a "compressed description" of the proteins or phenotypes offers a much stronger and much more direct target for analysis (even if this would introduce some further distortions, e.g., making it more difficult to account for regulatory sequences in DNA). It is harder for the proponents of causal parity to argue for "information content parity" e.g., by claiming that the environment somehow "describes" the organism in question (although one could try via the notion of affordances, see Gibson 2017; Gibson 1986; this would be a relevant concept also in the context of niche construction, see Pigliucci and Müller 2010a). But the proponents of causal parity could also use the AIT lens to strengthen their position. The concept of the genome "compressing" the organism requires a particular understanding of the developmental processes, one that the causal parity thesis is clearly geared against. Focusing on this point, highlights the discrepancies between interactionist (Kitcher 2003) and gene-reductionist (both deterministic and more nuanced) viewpoints.

While greater attention to the actual structure and concepts of information theory could have aided in those debates (which now seem to have died down significantly), in the current context, the crucial benefit comes from the fact that centring the information transmission vs. content distinction allows for distinguishing the instances where information theory is relevant for understanding BIOLOGICAL COMMUNICATION as a specific causal process (see also the discussions regarding natural signs, section subsection 3.5.3).

4.2.1 *Biological communication as information transmission*

To recap from the previous chapter: the proposed concept of communication identifies three central components of the causal organization:

(1) transfer of a *certain mark* K,
(2) involvement of *distinguishable* entities, X (sender) and Y (receiver), and
(3) the presence of a *structure* S enabling such transfer.

All these components can be mapped onto the Shannon model of communication: the sender and the receiver correspond to the joint operation of, respectively, information source with the transmitter, and destination with the receiver. The mark is implemented as the signal, and the structure includes the information channel. This has a straightforward reading from an instrumentalist viewpoint. If a biological process can be appropriately characterized in terms of biological communication, elements of the Shannon model can—in principle—be identified, and the entry criteria for using the information-theoretic formalism are met. Indeed, the empirical data seem to agree with this assumption.

In the current chapter, I re-analyze the results of the study in the previous chapter 3 to highlight the use of the term "information" in relation to BIOLOGICAL COMMUNICATION (see a selection of those excerpts in box 4.1). This additional analysis has been conducted with standard keyword search for "information" and with an additional Word Sketch analysis focusing on the terms related to information theory (Step 4. Word Sketch of the methodology outlined in subsection 2.5.2).

Box 4.1: Examples of mentions of information selected from the results of semantic search discussed in chapter 3.

Examples where picked through qualitative assessment based on a keyword search, from the results with the best similarity scores in the study described in the previous chapter. Queries and similarity scores are omitted here, as they are not directly relevant to the current discussion (full results are available in the online supplement at: https://doi.org/10.17605/osf.io/ax4jm). References are placed in bracket, they include the corpus ID and sentence number. Some examples shown here are embedded in the context of natural signs and as such constitute instances of information flow rather than transmission, see subsection 4.2.4.

1. "Song is a sexually selected trait that plays a key role in mate attraction and territory defence in birds (Catchpole and Slater 2008). It is a complex and flexible signal that often conveys information related to different male characteristics, such as quality, condition or current motivation (Gil and Gahr 2002)." (395482, 14)

2. "Nevertheless, signaling is important because it constitutes a source of (potential) information for the recipient and a relatively cheap way for the sender to aim to influence the behavior of others (Fischer 2011)." (645745, 97)

3. "All animals, including humans, must have the capacity to acquire information quickly and store it for long periods of time. Animal signals, like language in humans, are a highly compressed form of information transmission. This requires synaptic plasticity that,

along with natural selection, is subject to built-in mechanisms for mutation or variation. In this manner, for example, the information content moving from one brain to another is constant yet the fine structure of the receiving organ is different from the sender so the array of synaptic modifications will arrive with new associations and contexts, preserving as well as altering that unitary piece of information [77,78]." (1462585, 163)

4. "For example, it has been shown that the preputial glands of mice and rats are sources of pheromones and secrete information-coding volatiles that vary in quality or quantity with gender, reproductive condition, and between individuals (Zhang et al., 2007a; Zhang et al., 2008a.)" (1633453, 10)
5. "The importance of redundancy in allowing multimodal signals to overcome noise leads to the prediction that multimodal signals might also be different from unimodal multicomponent signals in the tendency for the different components to carry redundant or nonredundant information." (1855883, 55)
6. "Although vision and tactile [sic!] have been suggested to be very important in social interactions for mediating the transfer of information, the full understanding on the ways these two sensory cues are used in agonistic communication remains unclear." (1926553, 28)
7. "Both maternal and direct environmental effects require the presence of developmental plasticity, in which the organism can develop into different phenotypes, depending on environmental cues, despite the presence of the same genes. Under some conditions, detailed below, parental effects provide the parents with a tool of transferring information to their offspring in a much more flexible way than their genes can do." (3166187, 9)
8. "Social learning includes a wide range of mechanisms through which individuals receive and integrate information from other members of their social group." (3223637, 303)
9. "The now classic Shannon-Weaver model (1949) aimed to describe the transmission of a message and the influence of potential distortions such as 'noise' (also described as the level of uncertainty) on that message. Although very successful in optimising the process of information transmission, the model has been criticised for describing communication as a linear process, solely running from speaker to listener, rather than an interactive process between a speaker and a listener, for omitting contextual environmental factors, and for ignoring the relational social aspects of communication between two people (Kincaid, 1979)." (246277387, 119)
10. "In a final step, we investigated the interplay between spindles and ripple-mediated information transfer. Ripples during hippocampal replay have been proposed to reflect information package transfer [2,3]. To quantify information transfer, we calculated time-resolved mutual information (MI) [37] between the EEG and MTL relative to MTL ripple events […]" (199491769, 187)
11. "Brain-to-brain coupling is mediated by sounds and bodily movements generated by the sender and picked up by the receiver. Thus, during effective interaction, the sender's brain actively and systematically influences the receiver's neural response. Lack of sender-receiver coupling indicates that information has not been successfully communicated." (8709249, 65)
12. "The brain lacks the anatomy to support independent connections from all sources of evidence to all possible intentions-that is, the circuits that represent them. Instead the communication must share connections, and this invites some form of time-slice multiplexing. It is not possible for every source of evidence to communicate with the circuits that form decisions at the same time." (224817650, 596)

13. “These include signalling that an act is meant to communicate and that it is recognized as such, and that it can be deciphered for what it is supposed to mean—not just in terms of the semantic ’bits’ of information encoded but in terms of what the speaker intends to achieve with communicating them at a particular moment in time (i.e. the pragmatic meaning and illocutionary force needs to be derived).” (251020763, 13)
14. “Intracellular signaling networks in a cell respond to incoming signals to regulate some target molecules, to properly control the cell function. In general, signaling networks have multiple inputs and multiple outputs. The inputs can be ligands that, upon binding to their receptors on the cell membrane, create a chain of interactions through some intermediate signaling molecules, such as receptors, kinases, phosphatases, etc. This way the network outputs, typically target proteins such as transcription factors, are collectively regulated to produce an appropriate response. One possible way to model a signaling network is to consider it as a communication channel[1,2].” (202554259, 11)
15. “The released transmitters can then interact with one or a few adjacent cells to allow for information passage within a neural network.” (267632567, 18)
16. “This is a reasonable model for a network that just transfers the information from its inputs to its outputs, without any processing or computation on the information. This is typically the case in man-made communication channels[3]. In such systems, the output and input are ideally the same, if there is no transmission error[3]. In contrast, in signaling networks, outputs are typically computed from inputs, and there is a desired function that maps the inputs to the outputs (examples are provided later in the paper).” (202554259, 35)
17. “Signaling pathways at steady-state appear to encode only enough information to distinguish between two types of environment, but the information substantially increases if the downstream biochemistry can sense the response’s dynamics. Encoding different extracellular signals in the dynamics of signaling molecules can also coordinate downstream responses.” (231679501, 180)
18. “Signaling pathways that share common components require mechanisms to ensure specificity of information transfer, and this specificity is often provided by adaptor molecules.” (58547879, 167)
19. “Basically, when a ’sender’ module transmits information to a ’receiver’ module, its (dynamic) state is changed by the physical mechanism that allows connection to the receiver module. The extent of this change depends on the physical characteristics of the interconnection and increases with increased ’flow of matter’ between the sender and the receiver.” (19640778, 16)
20. “Electrical signals are transmitted back towards the inner retinal layers by bipolar cells, horizontal cells, and amacrine cells in the INL, and conveyed to RGCs in the GCL. Finally, this information is carried by RGCs to the brain.” (248153484, 32)
21. “Wiring transmission (WT) can be defined as a point-to-point communication, relying on the presence of well-defined structures through which the signals are transmitted, as virtual wires connecting two specific elements (Figure 1B). Chemical synapses represent a prototype for this kind of communication. […] Besides chemical synapses, also gap junctions, or electrical synapses, constitute an integral part of WT, which can be found in the mammalian CNS. The functional properties of WT are closely tied to the structural characteristics of synapses. It is a fast and highly specific communication modality in which the source of the signal and its target are in a 1:1 ratio. The presence of physically defined structures ensures relative stability in the connection between the source of the signal and the target. Thanks to those features, WT appears to be particularly well suited

for prompting activation or inhibition of effector systems as well as for guaranteeing an oriented flux of information through those "hardwired" networks." (265211080, 37)

22. "This interaction does not occur in the presence of penicillin. This result suggests that binding of the antibiotic to BlaR1 might entail the release of the interaction between L2 and BlaR-CTD, causing a motion of the α-helix bundle and transfer of the information to the cytoplasm of the cell." (25526234, 3)

While the corpus contains mentions concerned directly with the formalism of information theory and its application as a framework for modelling some biological phenomena and (more frequently) analysing experimental data (e.g., "We will now quantify how many messages can be transmitted reliably in the presence of not only crosstalk, but also biochemical noise." (16061300, 228) or "[Network inference] methods frequently use measures of coexpression, such as correlation or mutual information, to predict interactions between, for instance, transcriptional regulators and their regulons." (14633605, 146)), the semantic search results picked out only some of them, further suggesting that modern information theory focuses more on information content, abstracting away from Shannon communication model. Instead, the mentions of "information" listed in box 4.1 have a distinctly non-mathematical character, while at the same time drawing strongly from Shannon's conceptualization of information. This is confirmed by the Word Sketch collocations of "information" in these results. The collocations pick out technical terms such as "mutual information", "information source", "information capacity", or "information games", as well as more general like "information transfer", "transmission" or "flow".

These uses of the framework seem to fit the characterization proposed by Arnon Levy (Levy 2011, discussed in detail in subsection 3.3.1): they identify information transfer as a particular causal pattern characterized by directionality, systematic covariation and the activity in the causal relationship between the sender and the receiver. In the examples, the senders and receivers are quite universally identified. In most cases, individual descriptions focus on just a single direction of the connection—even if we know that in reality, these are much more complex and intertwined relationships, with bidirectional connections (note that this connects to the "response sensitivity" condition discussed in subsection 3.5.3).

Indeed, in many cases, as the examples highlight, these connections can be abstracted as unidirectional since the different directions that are involved are biologically implemented via distinct mechanisms: in the way that dendrites and axons are distinguished elements of neurons, or receptor molecules differ from various mechanisms involved in sending intercellular signals (ligands, vesicles, etc.). This is however not a universal property, and even in cases where such distinctions exist, more exotic effects—like ephaptic coupling in the brain (see e.g., Anastassiou et al. 2011; K.-S. Han et al. 2018; Pinotsis and E. K. Miller 2023)—can complicate these default relationships.

Indeed, several of the examples above are not precise in identifying the components involved and either implicitly ("downstream biochemistry") or explicitly ("one or a few adjacent cells") allow for a significantly more complex structure compared with the Shannon model. This underscores the difficulties with clearly identifying what components of the communicative systems are senders or receivers. This problem is unique to the case of biological communication, and further exacerbated by the larger difficulties of defining and delineating biological individuals (see Clarke 2010).

However, these conceptual difficulties do not seem to impact the textual practice significantly: notions of information transfer or transmission are applied to the biological processes directly (i.e., without any epistemic hedging) even in cases where the identity of the purported sender is quite dubious as in the example 12 in box 4.1, where the authors note the "evidence source" as the sender, blurring the boundary between instances of natural signs and actual communication. Together with the broader issue of limited ergodicity of the processes in the biological realm, this means that two central assumptions of the Shannon information theory—central reference point whenever biologists and cognitive scientists invoke "information"—are not met for a wide gamut of biological processes.

There are two possible ways of viewing this tension: either through an instrumentalist or a realist lens.[10] If we take an antirealist stance and view "information" as a useful fiction, these limitations are not severe, as the anti-realists have no trouble with admitting and accounting for a widespread presence of idealizations in science. There is now also a number of different positions for advocates of scientific realism to account for idealization, as a major feature of the scientific practice (see Potochnik 2017), even though historically this has been posed as a challenge to some flavours of realist approach to scientific knowledge.

Having set the scene, I will turn now to how the assumptions of a sender/receiver model and ergodicity impact biological use of Shannon's framework. First, I will explore the instrumentalist view, a recurring lens on biological information in philosophical literature, and then turn to the alternative, realist reading. While the realist perspective in my view offers a better handle on the conceptual practice in this case, I do not believe it is possible to decisively reject the instrumentalist reading. Hence, in what follows, I'm exploring the motivations of both views, while the reader is welcome to pick their preferred alternative.

---

[10] A fascinating element of these examples that falls outside the scope of the current investigation is the apparent reification of information. Information is treated as something quite literally transmitted over the communication channel, essentially equated with the signal (rather than the more subtle interpretation stemming from Shannon framework which would posit it as the change in subjective probabilities held by the receiver/destination). This connects closely to the complexities of physical basis of information that I have signalled previously, see footnote 2, and further links to the considerations regarding the functional or material basis of communication important for the analysis in chapter 5. Unfortunately, a detailed account of those intricacies would require a separate, more metaphysically-oriented investigation and will have to wait.

4.2.2 *An instrumentalist model*

In some instances, information theory is explicitly used as a modelling framework for biological phenomena. Among the examples in box 4.1, the following sentence shows it most directly: "One possible way to model a signaling network is to consider it as a communication channel" (example 14). Such examples can be viewed through the lens of Levy's fictionalist account of information talk (Levy 2011) or in terms of instrumentally interpreted model-based science (see Frigg and Nguyen 2020; Humphreys 2002). What connects these two perspectives is that they cast these ascriptions of informational properties as serious uses of a particular formal description, rather than purely rhetoric devices or examples of conceptual imprudence on the part of the scientists. Nevertheless, they view the description as literally speaking false. In either of those views, the identification of the senders and receivers involved can be easily indexed to the description and the researcher, as it is not taken to track any biological reality.

For Levy, biological information talk is a useful fiction aimed to capture some characteristics of the studied phenomena (the distinct causal pattern). In the pretence theory of fiction he draws on (Walton 1993), these fictional descriptions are governed by a set of rules (called "principles of generation") which posit a particular logic for the pretence—a logic which depends in non-trivial ways on non-fictional facts (for a more recent review of fictionalist positions, see M. Suárez 2008).

To build on Levy's example (2011, pp. 648-649): the description that "There's a reindeer in the sky" when I'm playing make-believe that clouds are animals depends on there actually being a cloud in a shape that resembles a reindeer. The principles of generation act as success conditions for the fictitious statements. Depending on the shape of the cloud, other participants of the game can either confirm my statement, claim "No, it's actually a goose" (to which I may respond "Oh dear, you're right!") or by saying "Come on Wiktor, you don't have to play" if I continuously disregard the rules and there is no cloud at all where I'm pointing (or the cloud is *clearly* a dog, I mean, look at the tail; and I'm just being annoying).

Similarly, in the case of information talk: the claim that "the signalling network is a communication channel" may literally be false (if, for instance, our understanding of communication is firmly on the personal level and requires complex cognitive characteristics) but may constitute a coherent fiction which tracks some real facts about the process—the model of communication I postulated in the previous chapter or Levy's three properties (of directionality, covariation, and activity)—that would be difficult to capture otherwise (for instance, in a highly complex process). But while Levy succeeds in highlighting the epistemic benefits the information perspective brings through fictionalist lens, the general approach he proposes is slightly too generic to fully take advantage of information talk through instrumentalist lens.

An alternative comes from the model-based view of science. My charge of genericity is concerned with the fact that the fictionalist account disregards that information talk is embedded within a formal theory. Instead, if we take information transfer as a formal model of various biological processes (even if it is often not explicitly mathematized in its various applications), we can supplement the view that Levy offers. Paul Humphreys (2002; 2019) offered an account of computational models as based on "computational templates". Such templates include "differential equation types, such as Laplace's equation and the Lotka-Volterra equations; statistical models such as the Poisson process and its various extensions; and specifically computational models such as cellular automata and spin-glass models" (Humphreys 2002, p. S2). In his view, computational (or formal) templates are a characteristic, basic form of computational models which can be applied across a number of phenomena, and a variety of scientific disciplines. Crucially, they have only the mathematical interpretation, can be computationally tractable, and have a clear set of conditions for application. To produce a full model, the templates have to be complemented with an interpretation from the target domain[11].

The Lotka-Volterra model is perhaps the most famous in that regard (at least in philosophical treatments). Originally the model has been proposed independently by two researchers. Alfred Lotka advanced it as a theory of autocatalytic chemical reactions (Lotka 1910) and then extended to the interactions of plants and herbivorous animals (Lotka 1920). Vito Volterra developed the set of equations to describe the predator-prey fish population changes during the World War I in the Adriatic Sea (Volterra 1931). This model has been later extended to a host of economic phenomena (see e.g., Chao 2023). The Lotka-Volterra equations are two differential equations that connect the changes of two variables: the particular interpretation of these variables (e.g., as population densities of predator and prey species, or as "employment-labour ratio" and "labour share of output", see Chao 2023) is the interpretation necessary to transform the template into a model of a particular phenomenon.

The distinction between the "information theory" and the "mathematical theory of communication", as Shannon pitched his framework, could be cast precisely as a distinction between a computational template (the equations for Shannon entropy, channel capacity, etc.) and the computational model (the equations as interpreted in the context of an engineered communication process). The template has been developed in a particular context but turned out to be applicable far beyond that context (see Shannon [1956] 2021).

These applications can, but need not be motivated by the commonalities between the phenomena (Humphreys 2002): in this context we can view the causal pattern highlighted by Levy (the directional covariation between active senders and receivers) as a precondition of the application of information theory,

---

[11] This can be also quite clearly presented in terms of the DEKI (denotation, exemplification, keying-up, and imputation) account of scientific representation (Frigg and Nguyen 2020): the formal template is a set of model properties that can be interpreted to exemplify properties of a number of target systems, with the key (a mapping connecting model and target properties) defining the relevant interpretation for a particular use of the model, and as such being domain-specific.

which the application of the framework brings to the fore. But the template view further highlights the possibilities of trans-domain integration that information theory—as a formal template—enables and provides a clear way of interpreting the results of computer simulations of biological signalling networks (e.g., the gene regulatory networks' framework, see Karlebach and Shamir 2008) as how-possibly explanations (see Bokulich 2014).

For instance, the concept of "redundancy" and "noise" (as in example 5, box 4.1) are treated as formal (even if not mathematized) concepts from Shannon information theory that are interpreted through the lens of the domain theory into a formal model. In the example mentioned, redundancy is understood as the partial overlap of sensory information provided through different modalities—even if this particular use of "information" relates to instances of natural signs, see also subsection 4.2.4.

Modelling sensory neural pathways as communication channels provides epistemic benefits. It highlights that they may be structured differently to perform different functions from channels carrying "unimodal multicomponent" sensory signals (e.g., the "where" and "what" components of vision), as in example 5. The formal templates of information theory allow for simulating these scenarios and comparing them to actual empirical data: an aspect missing from Levy's fictionalism. In this context, Humphreys takes a "selective realist" perspective which views some formal concepts as instrumentalist, while allowing other representational apparatuses to be interpreted realistically. In the context of information theory, this could mean that some applications of the framework are concerned with literal, realistic ascriptions of informational properties (e.g., in the nervous system or in some instances of animal communication). Meanwhile, others (e.g., in gene regulatory networks, as in example 12 above) could be treated as literally false but useful, instrumental descriptions of the processes. It is an interesting possibility, especially if we take into account the results from chapter 3: if such realist/anti-realist distinctions do actually hold, they are not clearly visible on the textual level.

4.2.3 *A realist model*

Alternatively to this fictionalist approach, we can consider the statements involving information talk as motivated by a realist interpretation by scientists who use the framework. This reading has a higher a priori likelihood considering the emerging empirical evidence of natural scientists' preference towards scientific realism (Beebe and Dellsén 2020). Further, it responds better to the fact, as noted above, that epistemic hedging (the use of "scare quotes", phrases like "as if", etc.) is relatively rare when information theory is involved. Nevertheless, the realist interpretation needs to face the fact that the assumptions of information theory are widely incorrect with regard to biological systems. This concerns in particular the unclear status of senders and receivers, and the pervasive non-stationarity or at least non-ergodicity of biological processes that are presumed to involve information. I will consider these problems in turn: they are important concerns for a realist interpretation of information-transfer talk, as these are crucial assumptions of Shannon framework which are literally false (or have limited applicability) in the context of biological processes. As mentioned before, instrumentalists have an easier task in this context, as falsity is assumed in their model, and because of the benefits they still provide. Realists need a more careful response to their literal falsity, similarly to other types of idealizations (L. Nowak 1980; N. Cartwright 1983; Weisberg 2007; Potochnik 2017). Let's consider the two assumptions—the sender/receiver model and the ergodicity of the information source—their limitations in biological context, and how realists can deal with that.

Realist response to the sender/receiver model limitations

As mentioned in various places throughout the thesis, life sciences do not have a clear way of identifying biological individuals systematically. This is especially the case when the question of individuality plays out across various scales of biological organization as in the case of informational phenomena. The fact that senders and receivers can be individuated both and either/or spatially and temporally further complicates matters in this regard (e.g., when memory is considered as information transmission).

Consider the case of "information packet transfer" in the hippocampus (example 10 in box 4.1) where the relevant components are either distinguished spatially through a functional role they play in information transfer (leading to circularity) or temporally (leading to anti-realism, as such temporal distinctions are highly dependent on the actual measurement method employed). Now compare it with the case of "information-coding volatiles", such as pheromones (example 4 in box 4.1) which are secreted by preputial glands but send information between whole organisms (mice or rats; the glands effectively act as encoders, while the organisms are an information source). These two cases activate two sets of philosophical issues, with the latter mostly recognized in the debates of biological individuality (Clarke 2010), and the former related to the broader problem of defining "systemhood" (see Rosen 1986).

The state of the biological individuality debate is quite telling in this regard: at its peak, Ellen Clarke summarized the various positions advanced and came up with a list of no less than 13 "serious contenders" (Clarke 2010, p. 317) for differentiating individuals from groups, a list containing as distinct properties as "spatial boundaries/contiguity", "sex", or "histocompatibility". Crucially, as perhaps obvious to biologically-minded readers, this pluralism leads to inconsistent characterizations of many boundary cases, where different criteria delineate individuals differently. This complicates a realist pluralist approach to the question of individuality. One possible way is the perspectival realist interpretation recently suggested by several authors (e.g., J. Suárez and Stencel 2020; Cárdenas and Javier 2022; Veigl 2022).

Perspectival realism (Massimi 2022; see also Massimi and McCoy 2019) is an attempt to reconciliate the increasingly shared acceptance of the fact that science is a human practice, always embedded in a particular cultural, social, historical, economical context, which impacts the knowledge that science produces (e.g., Giere 2006), with the claims of scientific realism. The view is closely related to Sandra Mitchell's integrative pluralism (S. D. Mitchell 2002, 2003, 2009), which seeks to highlight the compatibility of alternative models (due to their differing assumptions, the focus on different timescales, etc.), and the integrative character of explanations. Further, it connects to the Wimsattian flavour of realism (Wimsatt 2007), which adopts a broadly pragmatist stance towards individual scientific theories and models, but views their point of convergence in realist terms.

The thread running through all these approaches is that the claims of realism are addressed towards phenomena that are "robust", to use Wimsatt's term, namely—points of intersection of different scientific perspectives (e.g., Massimi 2022, ch. 6). But given Giere's cognitive motivations—in defining the concept of a scientific perspective, he draws on the study of colour vision and colour as a prototypically perspectival property—the concept of perspectives can be generalized beyond the human (scientific) cognition. In the context of biological individuals, this direction has been suggested by Javier Suarez and Adrian Stencel.

They consider the example of the disputed individuality of the holobiont (Margulis 1991). A holobiont is the assemblage of a host and a number of accompanying species interconnected through obligatory symbiotic relationships, which can be either regarded as a single organism or as a network of connected organisms. Suarez and Stencel argue that "[h]ost and microbes differ so much that what constitutes a part of a biological individual from the perspective of the host, can perfectly be an independent biological individual from the perspective of the microbes" (J. Suárez and Stencel 2020, p. 14). The "perspective" here needs not be read strongly for this view to work (i.e., we do not have to presume any cognitive capacities on the part of the microbes invoked here, or any other biological system considered—though, as I will briefly discuss in chapter 6, we could if we wanted to). We can take the "perspective" to define a certain relevant spatiotemporal scale or a certain coarse-graining of phenomena (see Flack 2017, the concept is discussed at greater length in subsection 5.5.2) in a way that is grounded by facts about the phenomena, and independently of external observers' (scientists') viewpoint.

This understanding of a "perspective" allows for interpreting information transfer realistically even if external (scientists') perspectives are not able to clearly delineate the individuals—senders and receivers—involved. Information transfer becomes indexed to a particular characterization of the process. However, if the characterization is consistent, e.g., from the perspective of the sender, receiver, as well as the signal, we can take it as identifying phenomena that are at least somewhat robust and have some degree of reality.

To illustrate the benefits of this approach consider two of the examples from box 4.1. Example 10 discusses the mechanism of information transfer in the hippocampus, as involved in memory processes. Example 20 discusses the particular, physical channel for the transmission of information from sensory organs (retina in this case) to the relevant brain area.

In the former case, the information transfer is considered through time and the employed analysis methods centre on the putative signal: the "ripple", a sharp-wave pattern of electrical activity (~80-120Hz in humans) generated by the hippocampus and associated with reactivation of novel knowledge (e.g., Buzsáki 2015). The senders and receivers are not explicitly marked, instead they are posited as temporal coarse-grainings of the process, with the amount of information (measured as mutual information) being dependent on the particular timescale considered. In this context, I would argue that the study of the signal does not consistently pick out robust senders and receivers (these are spatially co-located with one another, and differ along the temporal direction but in a somewhat continuous manner). As such, it should not be interpreted realistically as an instance of communication (which does not mean that the attribution of information is not realistic—in this context, the information talk could be realistically read as operating within the framework of information flow, as discussed below, subsection 4.2.4).

This is different from Example 20, where we have a chain of instances of information transfer with the senders and receivers being clearly delineated (not only from the perspective of the particular communicative phenomenon considered, but also according to some criteria of biological individuality—the presence of cell boundaries, the cells' spatial contiguity, etc.). They are senders and receivers robustly, even if we consider the "layers" invoked in the example, a concept which in other contexts has come under some scrutiny regarding its accuracy as a relevant unit of brain organization (e.g., Haak and Beckmann 2020).

Thus, the notion of robustness and the broad view of what constitutes a relevant perspective identifies instances where the formal framework of information transfer can be accurately applied (and interpreted realistically). These can then be distinguished from other phenomena that are embedded in the alternative formalisms—or where the notion of "transfer" is used instrumentally.

Realist response to limitations of ergodicity

The problem of non-ergodicity of biological processes (see e.g., Kauffman 2019; Medaglia et al. 2011; Weron et al. 2017) poses a similar challenge: it is a strong assumption of the Shannon information theory that the information source can be described as a stationary, ergodic probability distribution (and relatedly, that the Markov condition holds)—a condition often not met in the biological realm. There's a range of arguments regarding the non-ergodicity of the biological realm.

Kauffman's (2019, pp. 2-3) combinatorial argument states that the there's more possible combinations of chemical elements than (the organization of) physical systems could possibly explore in the lifetime of the universe. Hence, the universe above the scale of all possible stable atoms is non-ergodic. In the case of life, Kauffman points out that a short protein of 200 amino acids has already more possible combinations of existing amino acids than could be explored (with 20 amino acids, there's approximately 10^{260} possible combinations for such a sequence). While this argument might well be correct, it does not directly apply to the case of biological communication. While the biological state space is non-ergodic, it doesn't directly follow that *within* that state space the biological systems do not behave ergodically.

Then again, even if a biological process is non-ergodic (as often is the case, in a stricter sense than Kauffman considered, e.g., Medaglia et al. (2011); Weron et al. (2017)), in the context of communication it might behave ergodically (i.e., act as an ergodic information source). This might be the case for relatively simple communication systems, e.g., in vervet monkeys, because of their limited repertoire of calls: under the standard analysis, which identifies three separate call-types, we can likely treat the system as ergodic (though the argument possibly breaks down when we consider the variation in individual call-tokens, see Dubreuil et al. 2025; Price et al. 2015). All the three possible states will be eventually visited, and this is what allowed for the analysis of the calls in the first place.

Nevertheless, this is a relatively rare feature of biological communication. If we think about two other essential examples recurring throughout this thesis, the case of natural languages and the case of signalling during organism development, we can see that in neither of those two cases the information source is ergodic.

For natural languages, there is a formal proof to this effect (Dębowski 2021). The motivation for Dębowski's argument departs from the linguistic facts underpinning topic modelling (see chapter 2). Relative word frequencies, which we use to model the topic of a text, differ significantly between texts. As there is no infinitely long text (and indeed any natural language production will be relatively short), there is no reason to expect these relative frequencies to converge at some limit (Dębowski 2021, section 1.11 and 8.3).

For developmental systems I can only hand-wave: their adaptive character (see e.g., M. Levin 2021) directly connected with the structure of the signals (perhaps even implemented through signalling) means that they cannot be ergodic in the relevant scope (though possibly they're asymptotically ergodic, see below). This is because the underlying probability distribution of signals (information source) fluctuates in the course of development, regardless of the limited repertoire of possible messages. The information source needs to adapt to varying environmental conditions, to failures in course of the development, etc. Hence, even if there are some biological communication systems that meet the requirement of ergodicity, this is by far not a universal property.

While the anti-realists can simply treat the ergodicity assumption as false and disregard it, the realists have to provide a more detailed story. The assumption of ergodicity is a Galilean idealization (L. Nowak 1980; Weisberg 2007): it introduces "distortion into theories with the goal of simplifying theories in order to make them computationally tractable" (Weisberg 2007).

Shannon introduces this assumption as the mathematics required to deal with Markovian, ergodic stochastic processes are much simpler than the ones required for non-ergodic phenomena. Indeed, these methods are still largely lacking, even several decades after the mathematical theory of communication was introduced. Hence, ergodicity ensures computational tractability of Shannon framework. As Weaver writes (Shannon and Weaver 1964, p. 12), "[e]rgodic systems […] exhibit a particularly safe and comforting sort of statistical regularity." Shannon relies on this "comforting regularity" for multiple elements of his framework: for instance, inferring the entropy of an information source (or any other statistic of the source) directly from the statistics of observed messages (an important fit for cryptanalysts) is only possible if we can assume statistical homogeneity of the source.

Further, the ergodic source can be considered as operating (in the limit) at the statistical equilibrium, i.e., to be accurately described by its descriptive statistics (such as mean and variance), without considering the initial conditions or its history (additionally ensured by the Markov condition). This is crucial for the computational tractability of the communication processes, and for considering efficient or optimal coding algorithms.

But it doesn't mean that non-ergodic processes cannot be placed within the scheme of information transfer. One example is a special class of asymptotically ergodic processes[12]. Some propose to consider learning processes in those terms (assuming that there's a final stage of "complete" knowledge), and if we take into account the recent attempts at bringing both learning and evolution together under the umbrella of Bayesian inference (e.g., Csillag et al. 2025; Czégel, Zachar, and Szathmáry 2019; Czégel et al. 2021), this kind of argument could be extended into the domain of systems undergoing natural selection (with the limit corresponding to a "perfect adaptation" to a given niche—an admittedly hazy reference point).

The Galilean idealizations are considered as easily reconciliable with scientific realism (Weisberg 2007). Leszek Nowak describes the scientific method which builds on such idealizations as one which proceeds from the abstract to the concrete (L. Nowak 1980, p. 95). In his discussion of scientific methodology, Nowak proposes a realistic lens on idealization. At the same time, he does not fall into the view that idealizations, and other types of systematic falsities, are a step in attaining better, more factual, theories (a view inconsistent with actual scientific practice, where such inaccuracies are persistent and not considered as inherently problematic, see Potochnik 2017).

---

[12] I'm grateful to Łukasz Dębowski for pointing this out to me in a conversation at the 11th Peripatetic Conference on Cognitive Systems Modelling, Zakopane, October 2022.

Nowak's Marxist reading of science focuses on the consequents of idealized laws,[13] which can "represent actual regularities" (L. Nowak 1980, p. 133): the step from abstractness to concreteness is one from more generalized patterns indicated by idealizations, which are incorrect, but nonetheless capture something about the actual regularities. This knowledge can be used wherever required. For example, the idealization of the movement of a ball along a frictionless plane gives us directly some scientific knowledge, but can further guide the more "concrete", less idealized scientific practice, for example in an engineering context.

In the case of ergodicity, this model of thinking about idealizations suggests that the "optimal" reference point captured under the assumption of an ergodic (and Markovian) information source explains something about the non-ideal (non-ergodic) processes involved in information transfer. Especially since this idealized, ergodic process can be closely approximated in the engineering context, it acts as a limiting case for the non-ergodic system, providing a boundary on what is possible in the non-ergodic scenario. Limiting cases can be considered an important element of exploratory practices of science (e.g., Shech and Gelfert 2016).

Alternatively, we can assume that the behaviour of the system is monotonic with regard to the ergodic cas as the limit. In such case, studying the limiting, ergodic system and considering the distance of the actual, non-ergodic behaviour from that limit can have genuine explanatory value (Laymon 1991). However, this assumption of monotonicity is quite strong and does not hold in regard to non-ergodic systems in general, except for the asymptotically ergodic systems. In asymptotically ergodic systems, the entropy rate (a function assigning entropy to a stochastic process) equivalent for a non-ergodic information source, the asymptotic rate, is relatively smaller, as in the limit it approximates the maximum entropy rate of the ergodic component of the information source (see Winkelbauer 1970)—following the kind of reasoning suggested by Ronald Laymon.

This tension surrounding ergodicity may be an additional explanation for the strong preference of researchers (in the animal signalling case) to include natural selection or adaptation in their models of biological communication, as the evolutionary pressure can ensure that the process considered can be meaningfully modelled as (at least asymptotically) ergodic. In this sense, if one wants to adopt the Shannonian model of information transmission to model biological communication in general, the model introduces important constraints onto what kinds of structures (in the sense of the definition in chapter 3) can actually enable communication.

---

[13] Nowak focuses on laws, as the dominant view of the time was the deductive-nomological account of explanations. However, he explicitly mentions that the "the axiomatic method, the method of modelling, of approximation, etc." are to be derived from this central model of idealized laws. Hence, we may take "laws" here as a stand-in for a more general type of "scientific cognition".

4.2.4 *Information flow and natural signs*

As noted previously (in subsection 3.5.3), the definition of communication, if too liberal, leads to mistaking genuine instances of biological communication with the phenomena belonging to the broader category of natural signs. This is paralleled by how the formal information transmission framework can easily be misapplied to cover processes from the broader category of information flow. Indeed, some results discussed in subsection 3.5.3 do include references to information as not belonging to a transmission schema, and instead e.g., picked up from the environment. Consider the examples in box 4.2.

Box 4.2: Examples of mentions of information in the context of natural signs.

These examples do not fit into the Shannonian information transfer framework, but in some cases can be formalized in terms of information flow. Examples were selected from the occurrences of the word "information" in top examples (highest similarity score to queries) produced by semantic search results from chapter 3 through a close reading, interpretative procedure.

1. "Finally, I examine a wholly different class of information, one that pertains to the nature of the individual. I call this information "indexical," and suggest that human and non-human primates are remarkably similar when it comes to this class of inter-individual communication. Here I distinguish between information that is sent in the form of signals from information that is emitted in the form of cues." (233236380, 37)
2. "The importance of available tactile information in new terrain is suggested by studies that manipulated the antennae." (155626, 47)
3. "The question of whether human smiles convey information about both dominance and prestige is further complicated by the realization that human smiles have several functionally distinct forms that range from happiness smiles to embarrassment displays." (19584746, 42)
4. "This synchronisation of rhythms is referred to as entrainment, and facilitates optimal sampling in discrete time windows. It allows the information in the signal to be divided into meaningful chunks, which can then be processed and understood. Subsequently, this information may be used to predict the upcoming signal, whereby higher order regions may provide top-down feedback to modulate the entrainment of oscillations to the signal (Park, Ince, Thut, Gross, & Schyns, 2015)." (53338573, 16)
5. "Information about the position of sensory objects and identifying their concurrent behavioral relevance is vital to navigate the environment. In the auditory system, spatial information is computed in the brain based on the position of the sound source relative to the observer and thus assumed to be egocentric throughout the auditory pathway." (231838884, 0)

Example 1 makes the distinction explicit: there is a difference between "signals" and "cues", but both can be accounted for in terms of "information". In the case of "cues" it doesn't make sense to distinguish the sender—as discussed in subsection 3.5.3—even if there are two distinguishable entities involved, only the receiver can be meaningfully called so. Hence, the Shannonian information transmission model does not apply in those instances. Since the "source" is often an object, as in examples 2 and 4, or simply—it is not a part of the proper structure between the entities involved, these are not cases of information transmission. Nevertheless, there can be regularities—"local connections", in the sense of Barwise and Seligman (1997)—between the source of the cues and the receiver. In the case of cognitive systems, we may consider the receiver as "designed" in the course of evolution to perform that function (i.e., of receiving information), in which case it makes sense to describe the regularities in terms of infomorphisms. This is the case in example 2: through experimental intervention, the researchers modify the classification of the receiver so that the previous infomorphism doesn't hold. Effectively, information is lost as a result of (some) such interventions.

Example 4 is particularly interesting, as it exemplifies a much broader class of uses of information talk (see also example 10 from box 4.1) in the context of neuroscience. The example abstracts away from the scheme of information transfer, but it doesn't presuppose that one wouldn't apply. Consider the "rhythms" mentioned in this example as oscillations of electrical activity at the scale of several thousands (or more) of neurons. Then, if we take the perspectival route suggested above, the "signal" mentioned could allow us to meaningfully delineate some neural structures involved in the particular cognitive process in question. But this is not the point in the example: the notion of "information" refers to the internal structure of the signal. The lack of "sender" and "receiver" is not only a matter of reduced spatial resolution of the research methods involved (as "rhythms" are most often studied through EEG which pools together electrical activity from relatively large areas of the cortex). The information in the signal can be analysed without a reference to the source and destination, directly through the dynamics of the signal.

This is a special case of information flow: the infomorphism is assumed, but needs not be defined. What matters is the classification as carried by the signal, encoded in the patterns of its dynamic and how it unfolds over time. At the same time, the relevant informorphisms[14] are not only external to the "signal" or "cue", but also intra-systemic: they are correspondences between signals. This "systemicity" is an important property of some information flow structures and I will cover it in more detail in subsection 5.5.1.

The emerging picture casts information transfer as a special case of the broader class of information flow (similarly to how communication is a special case of natural signs). However, there are some processes that could be accurately described in terms of information transfer, but the formalism of information flow provides a better purchase on their relevant properties. Those processes require strongly structured signals and are often associated with the focus on intra-systemic infomorphisms.

---

[14] I.e., mappings between classifications—as we will see in more detail in chapter 5, this "encodingist" perspective is still dominant in neuroscience. Even if the correspondence relationships are not described in terms of infomorphisms, they can usually be modelled as such.

As I will cover in greater detail in chapter 5, the presence of strong intrasystemic connections between signs (or signals) is a required feature of some “more complex” forms of Peircean signs, symbols in particular. What the analysis of the example 4 above suggests, is that while such “systemic” signs emerge in communication-like process (amenable to analysis in terms of information transfer), they become a subject of study in their own right, outside the communicative structure in which they are embedded (i.e., through the notion of information content or information flow).

4.3 WHAT DO WE NEED "INFORMATION" FOR?

Shannon's mathematical theory of communication has brought the word "information" to the front of contemporary science, with its role for biology and cognitive science being perhaps the best example of its contemporary standing. Nevertheless, despite reformatting the debates surrounding biological communication, information theory has received mixed reception in this—so it would seem—natural context. This has been partially a result of some misunderstandings of the structure of information theory and its various flavours, as I've tried to show in the first part of the chapter.

Shannon's framework introduced some crucial assumptions which lost relevance in the face of further developments of information theory. These nuances, however, were overlooked when the mathematics of information have been adopted outside its original engineering context. Shannon himself noted that early on in the "bandwagon" paper (Shannon [1956] 2021), though the actual relevance of formal theory of information is in fact broader than he argued in that piece.

The complexities I have attempted to track in this chapter highlight that the concept of "information" at play is not a unitary one. In fact, it resembles a patchwork concept, as discussed at length previously in chapter 3. The different patches emerge from different directions the original Shannon's formalism took: the different formal notions of communication are associated with different modelling tools they offer—different techniques, in the sense of patchwork concepts. As INFORMATION reappears in various contexts, it adopts different roles and needs to be understood through different formal definitions. How to connect this pluralism about information with the monism with regard to communication I have advocated in the previous chapter?

The answer is likely obvious from the foregoing discussion: only the concept of information transmission emerging from the mathematical theory of communication bears directly on the questions related to biological communication. The remaining patches reflect the various uses of "information" that do not bear on our understanding of BIOLOGICAL COMMUNICATION. These other informational notions do have some crucial roles in biology. If the questions regarding the physical basis of information are taken into account, they likely converge along some dimensions.

However, it is the information understood in terms of Shannon entropy, as the amount of change in (subjective) probabilities at the destination, resulting from the reception of the signal, that is relevant to biological communication. As I have already argued in chapter 1, information is influence, precisely because a signal carries information in virtue of the change in the receiving system it leads to (contra Scott-Phillips 2008). In this respect the characterization of information theory proposed by Arnon Levy is correct in that what it picks out across different contexts—inasmuch as it is actually "information transfer talk", not just "information talk"—is the causal structure of that influence. But this causal structure is a prerequisite for the application of information theory to make sense. The epistemic payoff lies somewhere else.

The crucial benefit provided by information theory can be best seen through focusing on the signalling games framework in its modern version (see Skyrms 2010; O'Connor 2020). Signalling games are particularly interesting in this context as they build on the applications of the formalism of information transmission, sharing its assumptions. The relatively abstract description of a biological process that information talk affords—regardless if we're instrumentalist or realist—allows a grasp of the coarse-grained dynamics of that process, and together with the associated formalism, this enables research questions that otherwise would be outside the scope of biological and cognitive research.

The most obvious examples come from the study of the evolution of language as a particular form of animal signalling. In this context, a variety of features of real biological communication systems can be examined in the abstract to better understand how they contribute to linguistic phenomena. This is done in the studies of the origins of compositionality, i.e., the property of (some) linguistic expressions in virtue of which the meanings of tokens contained can be combined into the meaning of the expression as a whole. This line of research explores how a variety of factors supports the emergence of a compositional communication system: from the structure of sensory inputs (Lazaridou et al. 2018), via the size of vocabulary (M. A. Nowak, Plotkin, and Jansen 2000; Mordatch and Abbeel 2018), to the presence of noise in the communication channel (Kottur et al. 2017; Kuciński et al. 2021; for a broader discussion of this area of research see Rorot and Rączaszek-Leonardi 2023; and Zubek, Korbak, and Rączaszek-Leonardi 2024). But the application of signalling games is much broader: it includes cancer biology (e.g., Archetti and Pienta 2019; Bayer et al. 2022), neuroscience (Fenton et al. 2023) or the dynamics of mimicry (Casey, Massey, and Mishra 2021), effectively covering the range of scales where COMMUNICATION is attributed.

There is little explicit work connecting these various uses of signalling games. Philosophers focus on distinctly philosophical questions (e.g., Skyrms 2010). Cailin O'Connor (2020), for instance, focuses on methodological challenges that reappear across different contexts of use of signalling games (e.g., the focus on evolutionary stable strategies emerging from games, rather than on the dynamics of evolution more broadly).

Instead, biologists provide some important insights. In a pair of reviews covering cellular (Hummert et al. 2014) and molecular (Bohl et al. 2014) uses of evolutionary game theory (including signalling games), the researchers conclude that "[b]eyond providing a better understanding of some basic principles and processes of life, evolutionary game theory is also a promising tool for practical applications in medicine and biotechnology, producing counter-intuitive results where conventional approaches fail" (Bohl et al. 2014, p. 3073). The "basic principles and processes of life" in this context refer to the dynamics of evolutionary processes, which can be directly modelled in this approach. The "counter-intuitive results", on the other hand, include the games which lead to the observations such as the "survival of the weakest" (Frean and Abraham 2001; discussed in Hummert et al. 2014), an apparent contradiction to the standard Darwinist principle of "survival of the fittest".

This conclusion highlights that signalling games may serve two main goals: integrative and exploratory. Integration here is meant in the sense of integrative coordination (Neurath 1937; Potochnik 2011; I borrow the term from Miłkowski and Hohol 2021), rather than integrative pluralism (as discussed above in subsection 4.2.3), as what we consider are models from a single modelling framework, sharing many assumptions, and applied across different domains within biology and cognitive sciences. I opt for "integration" here rather than "unification" to avoid the reductivist associations of the latter term (e.g., Oppenheim and Putnam 1958).

The sort of integration at play can be viewed through the lens of knowledge transfer through formal templates (Humphreys 2002, 2019, the discussion from subsection 4.2.1 applies to signalling games as well). In considering the transfer of models across scientific disciplines, Humphreys (2019) focuses on unification and suggests that what may be attained in this way is a unified representation, rather than explanation. But the case considered here is one of a more local transfer, and—as I've underscored already—the uses of the template follow similar assumptions, enabling a stronger integration. But is it an integrated explanation?

Indeed, if we adopt the prominent new mechanist account of explanation (Bechtel and Bollhagen 2021; Craver 2007; Glennan 2017; Machamer, Darden, and Craver 2000) we can consider signalling games as at least attempting at an explanation of the phenomena, by identifying components, their relationships and activities. These are rarely full mechanist descriptions of the relevant biological processes, and indeed they have a more representational character, as the ultimate outcome of signalling games account of various biological (also social) processes is to arrive at an optimality model of the process. This is the purpose of identifying various equilibria in the analysis of the games—criticized in the Bohl and Hummert's pair of reviews (Bohl et al. 2014; Hummert et al. 2014).

Nevertheless, to achieve this optimality explanation, a signalling games model needs to explain at least some causal patterns at play. Importantly, a significant part of the explanatory power in those instances comes directly from the application of information theory required for the construction of a signalling game model (in the contemporary approach which ties the framework to information transmission). In this sense the uses of information transmission might be read through the lens of mechanistic explanation. Shannon's information theory in this context, because of its formal basis, offers a more powerful explanation than simply the application of a conceptual communication framework.

In terms of the exploratory power of signalling games, they allow singling out factors which in biological (cognitive) systems are deeply intertwined and cannot be manipulated directly. More broadly—the use of signalling game models allows for exploring the dynamics beyond the limitations of actual biological systems, including the boundary cases which may not be attainable biologically (see the discussion of ergodicity above). This means that on some occasions the signalling games models do not have actual biological targets. Those instances constitute cases of hypothetical or targetless modelling (in the sense of Weisberg 2013; see also Frigg and Nguyen 2020; Pero 2024). They provide counterfactual knowledge about the actual biological systems. Alternatively, if we focus on the fact that the models generalize biological phenomena, as in the case of the mentioned models of evolution of language, we could view them as providing how-possibly explanations for some observations.

These generalist conclusions do partially account for the role that information theory plays in modelling biological phenomena, as exemplified in signalling games models. But the formal framework plays also a more specific role: in abstracting away the material implementations of individual instantiations of biological communication, it enables a functional description and analysis of those processes. Beyond the integrative functions of such an abstracted description, information theory highlights common causal patterns in the phenomena, motivating theory development in life sciences and uncovering novel experimental directions.

4.4 SUMMARY

The purpose of this chapter was to explore the relationship between the emerging account of BIOLOGICAL COMMUNICATION proposed in this thesis and the formal framework of information theory, underpinning references to "communication" and "information" in biological literature. I have provided a brief historical outlook of the developments information theory has undergone since its inception and identified four somewhat separate formal frameworks: original Shannon's theory of information transmission, algorithmic information theory account of information content, Barwise and Seligman's theory of information flow, as well as informational signalling games. These approaches differ in terms of assumptions and conceptualizations of central terms, including "information." Using this view, I have then gone back to the corpus results from previous chapter to trace how "information" is used in communicative contexts.

The emerging view is that the notion of BIOLOGICAL COMMUNICATION can be formalized in terms of information transmission, but not content or flow. Even in that case, the formal framework requires careful investigation into the assumptions it brings that are often not met by biological systems. Here I have focused on two: the assumption of a sender/receiver model and of ergodicity of the information source. These assumptions are necessary idealizations for the application of the mathematics of information transmission.

I have argued that there is no available, general and non-arbitrary way to distinguish biological individuals—and resultingly, no general and non-arbitrary way to identify the senders and receivers. Nevertheless, in many of the relevant cases an internal (system-level) perspective on the communicative process allows to identify the relevant components in a robust way. In turn, the availability of such robust distinction can motivate a communicative description.

For ergodicity, while biological systems often lack it in the relevant scope, the admission of an adaptive—evolutionary or learning—context allows for ignoring the non-ergodicity of many cases of biological communication. In those cases ergodicity acts as a limiting case, allowing for an application of idealized model and inference to non-idealized cases. There is, however, a number of examples to be expected where even this asymptotic perspective fails. In such cases it is likely that the ascriptions of information should be read in an instrumentalist perspective—or perhaps, a realist application of a non-ergodic concept of information yet to be developped.

Nevertheless, contemporary approaches to scientific realism allow for some falisities in our scientific models, and hence the presence of those idealizations does not prevent us into reading most of the ascriptions of information in biological and cognitive literature from a realist perspective.

5

SIGNALS AND MEANINGS. SEMANTICS AT THE SCALE OF MOLECULES AND AT THE SCALE OF WORDS

Consider the following sentences:

> "Both systems have been extensively used for Ca2+-signaling research, but the meaning of their Ca2+ signatures remains elusive." (51934500, 51)
>
> "In humans and honey bees, the abilities of the listener directly constrain the meaning of information exchanged during an interaction." (227241123, 108)

This type of parlance, attributing "meaning" to a variety of biological signals, is not overly frequent in the literature, and often provokes suspicion in researchers. Nevertheless, it remains a regular feature of the literature on biological communication across scales, as we will see in this chapter. Indeed, while some ascriptions of semantics are correctly suspected to be anthropomorphizing or purely metaphorical (e.g., the now commonly accepted notion of "crosstalk" in cellular signalling, which obviously involves no "talk", e.g., Vert and Chory (2011)), claiming that cellular signals never have any content would not be justified too.

Clearly, neural signals are said to "encode" stimuli or memories (e.g., A. K. Robinson, Quek, and Carlson 2023; cf. Brette 2019), while the bioelectric impulses between cells of an organism undergoing morphogenesis are said to "encode" the target morphology (e.g., Oviedo et al. 2010; Vandenberg, D. S. Adams, and M. Levin 2012). This prompts an important problem that in the context of the brain and mental states has generated a lot of scholarship: how, working within the naturalistic paradigm, we can account for the fact that some brain states are able to "refer" to some objects, or why expressions animals produce—whether they're linguistic or not—"mean" something. The emerging picture of COMMUNICATION I have proposed so far, as something that is widely shared across the biological realm, provides us with intriguing conceptual challenges and opportunities: having established the tentative continuity of various forms of signalling, we can inquire whether there is anything like "content" across various scales, and how ascriptions of semantics differ in various research fields.

In philosophy of mind, an area I will be drawing on heavily in this chapter, this is an important open field of research. The majority of researchers accept that in order to account for how our thoughts can be "about" something, we need to develop an account of how to connect brain states to objects outside the brain, without circularly invoking the cognitive notions that we're trying to explain. In biology more broadly, this problem has come up in the context of animal signalling, but not much beyond that.

In that context, for quite a long time (see Allen and M. Hauser 1992), scientists believed that the non-human animals' expressions—vocal, gestural, chemical, etc.—provide information only about the internal state of the receiver: emotional or motivational. Over time, however, it became increasingly clear that this "Emotion" view, as Allen and Hauser call it, is not sufficient to account for all forms of communication, and an "externalist" alternative has developed, which allowed for the signals to refer to external states or objects. This Emotion/ Reference debate (Allen and M. Hauser 1992) is important both philosophically and methodologically. If an animal vocalizes an `alarm-call`, we will study it differently if we believe that it refers to the sender's internal state (e.g., their fear) or to the environmental state (e.g., presence of a predator).

Similarly, if a philosopher wishes to trace the evolutionary origins of human language to non-human animal signalling systems, these two hypotheses will provide them with divergent sets of explanatory tools. While currently the Reference side of the debate seems to prevail, the mechanisms of establishing reference even in this much-trodden area remain far from clear. The associated complexities are visible, for instance, in Ruth Millikan's account of naturalizing semantics (that I will explore in greater detail below). She notes (see Millikan 1995) that for a particular signal or representation, the direction of fit can be either "descriptive" (where the representation is supposed to conform to the world; inventory of what is in the fridge or a mother hen calling to its chicks to signal presence of food in some spot) or "directive" (where the world is supposed to conform to the representation; a grocery list or a mother hen calling to its chicks to make them come to that spot). These two directions presume very different kinds of meanings, and ultimately Millikan suggests that the most primitive biological representations and signals are a complex mixture of the two: what she terms pushmi/pullyu representations (in the mother hen example—the call refers to both the presence of food (descriptive) and includes a command for the chicks (directive)).

But what Millikan considers, is the scale of whole, multicellular organisms communicating with one another (even if her model is in fact applicable more broadly, she relies on these higher-level examples). And as we saw in the previous chapters, there's a variety of "signals" and "messages" across all scales of biological organization. In many cases the researchers seem to claim that these signals are indeed "about" something, whether internal or external, directive or descriptive.

As mentioned above, philosophers consider the problem of biological meaning primarily in the context of philosophy of mind. They ask how the brain/mind states can be about something in the world: e.g., how neural signals encode visual properties of the world which then can get transformed into neural correlates of language. This question in its modern formulation goes back to Franz Brentano ([1874] 2015), who used the term "intentionality" to refer to this "aboutness" of mental states. But in the second half of the 20th century, these debates were transformed by philosophers' focus on solving the puzzle of intentionality in a "naturalistic" way, i.e., drawing only on properties that fit within the picture that our (current or future) best science of cognition provides.

In that context, the issue began to expand also into the domain of semantics of language and other signalling systems—which are believed to depend on the referential properties of the associated mental states. For this reason, to address the broader problem of meaning in biological communication across scales of biological life, I will begin with the considerations originally aimed at explaining the mental states.

In various places in the chapter, I will mention the meanings of "signals", "signs", "representations", among others. This will be largely driven by the preferred terminology of the theoretical view I consider at any given point, but ultimately the picture I wish to suggest applies to biological communication in the sense discussed in previous chapters. The scholarship in philosophy of mind relies quite heavily on the notion of "representation," while I ultimately wish to adapt these views to the context of BIOLOGICAL COMMUNICATION, where—as established previously—"signals" are the primary currency. For these reasons, in the majority of this chapter I will use the notion most directly related to the area currently covered, assuming that these notions (signals, signs, representations, mental states) are equivalent with respect to how their meaning is established—or at least that the naturalistic explanation of "aboutness" in those cases is structurally analogous. This assumption is in fact widely represented in the literature, especially in how the considerations regarding linguistic reference (or, like in the examples Millikan considers, non-linguistic animal communication) are applied to mental or neural states.

In the current chapter, I attempt to provide an empirically motivated conceptual framework that accounts for the practices of ascribing meaning in life and cognitive sciences. To this end, I will use the previously developed understanding of BIOLOGICAL COMMUNICATION and conduct a second analysis of the corpus, focusing specifically on how biologists and cognitive scientists ascribe semantics to the processes they study. The analysis follows the same methodological schema, as already discussed in chapter 2. In what follows, in section 5.1 and 5.2, I review the existing philosophical discussions of intentionality, which inform both the corpus study (the design of the appropriate queries), and the analysis of the results. The results are presented in section 5.3. They highlight that meaning is understood differently in the literature, depending on the scale. Hence, in section 5.4 I analyze the practices of ascribing semantics at the "organism" level, and in section 5.5 on the "cell" level. In section 5.6, I conclude by contrasting the different accounts and highlighting their commonalities.

5.1 FROM FUNCTION TO MEANING

A golden standard in the naturalization of content in biological systems—both the content of internal states, and of communicative signals—has become the biosemantic approach of Ruth Millikan (Millikan 1984, 2004). Millikan builds on a previous theory of informational teleosemantics advanced by Fred Dretske (Dretske 1981, 1986, 1995, see Schulte and Neander 2022 and subsection 1.3.2). Dretske relies on the idea of information-carrying representations (in the sense of mutual information, see chapter 4), and claims that the representation's content is established by what is its function to carry information about. Consider one of the central examples in this literature (introduced in Dretske 1986; see also Schulte and Neander 2022).

There is a number of water-living species of Gram-negative bacteria, mainly anaerobic or with preference towards lower-than-atmospheric concentrations of oxygen (microaerophilic), that use the Earth's magnetic field to guide their movement away from the water surface (where the concentration of oxygen is higher)—a process called magnetotaxis (hence, this polyphyletic group of bacteria is called magnetotactic bacteria). These organisms have evolved special organelles called magnetosomes, which contain magnetic crystals, primarily magnetite (iron oxide) or greigite (iron-sulfide compound), which allow them to align to the magnetic field.

Now, the standard philosophical story is the following: the magnetotactic bacteria living in the Northern Hemisphere are north-seeking, i.e., they have a preference to swim in the direction of magnetic north, which in their standard environment aligns them downward in the water, towards lower concentrations of oxygen. As this was why the magnetosomes and magnetotaxis were selected for—hence, their function—we can take the magnetosomes (and bacteria) to represent the direction of oxygen-free environments (magnetosomes carry mutual information about the direction of such environments and it is their function to do so). If we take those bacteria and put them in the Southern Hemisphere, their magnetosomes would misrepresent the direction of oxygen-free sediments, and lead them to swim upward—in accordance with their preference for magnetic north.

This illustrates a number of concerns to Dretske's original formulation (1981): most notably "the functional indeterminacy problem" and the "problem of distal content". It is equally plausible to classify the function of magnetosomes in those bacteria as pointing in the direction of low oxygen, as well as towards magnetic north. However, in the latter case, no misrepresentation in fact occurs. Indeed, this is the problem of distal content: for there to be selection with regard to the mechanism (magnetosomes) being useful at representing distal features (low oxygen environments), there must also be selection with regard to proximal features (magnetic field). In which case, we have no grounds of distinguishing between either type of contents (or function), while both can lead to very different conclusions (as seen in the case of misrepresentation in the Southern Hemisphere).[1]

While Dretske sought to find additional criteria that would aid his approach in dealing with the indeterminacy problems (e.g., considering the role of behavioural flexibility and learning for establishing distal content, Dretske 1986), these arising issues motivated Millikan's work, which shunned the assumption of correlation or mutual information between the signal and what it's referring to.

Millikan's account presupposes a distinction into a producer X and consumer Y within the representing system[2]. When we consider the representations passed between the producer and the consumer, the biosemantic approach assigns content to those representations that belong to an interrelated set of various representations that can be shared from X to Y. These representations in normal conditions map onto (usually external to the X-Y system) states of affairs, objects etc., and in virtue of that, contribute to Y's performance of its proper function (with the direction of the mapping depending on whether the signal is directive or descriptive).

Consider one of Millikan's favourite examples of honeybee waggle dance (*Apis mellifera*, though other species of the *Apis* family perform dances as well; e.g., Millikan 1989b): a famous instantiation of insect communication, bees upon returning to the hive perform a "dance" in which they move around in a figure-

---

[1] In fact, the story is premised on assumptions that turned out to be significantly simplified in the course of more recent empirical research. First, some researchers argue that "magnetotaxis" is not a case of taxis, as the mechanism of action is purely physical—the presence of magnetic crystals in magnetosomes forces the bacteria into alignment with the magnetic field, even for dead bacteria (Dusenbery 1996, p. 100). This differs from other types of taxis, which involve active sensing (at least in the form of a stimulus-response setup) of their environment by bacteria. E.g., in aerotaxis, which often accompanies magnetotaxis, the bacteria have some sort of memory to compare concentrations of oxygen between two time points, and if the concentration increases (for anaerobic bacteria), they stop swimming forward and tumble to switch direction.

However, a philosophical point could be raised that since the magnetotactic bacteria actively produce magnetite or greigite, it still is a form of taxis—though the active sensing in question happens on the developmental timeline. This is further supported by the fact that magnetotaxis is often used in tandem with aerotaxis. Second, while the majority of magnetotactic bacteria in the Northern Hemisphere are north-, and in the Southern south-seeking, as per Dretske's original example, it is now known that at the equator, north- and south-seeking bacteria are equally present, that there are Northern south-seeking bacteria, and that magnetotaxis frequent pairing with aerotaxis makes the the preference for magnetic north or south plastic, so that it can change or be lost if the bacteria are placed in an environment where there is no selective pressure (uniform oxygen concentration; see the review in Lefèvre and Bazylinski 2013).

This larger complexity could be used to disambiguate functional indeterminacy, though the larger point Dretske raises still stands—that the problem of indeterminacy might not have a *general* solution.

[2] There are different terminological preferences among the theoreticians to call the two sides involved in communication the producer or the sender, and the consumer or the recipient. The producer/consumer terminology is more popular among the proponents of teleosemantics, while sender/recipient more widespread in accounts embedded in information theory. While the two terminologies have different connotations, in what follows I will treat them as equivalent and switch between them depending on the more natural notion for the particular account considered at a given point.

of-eight movement split by a straight walk forward associated with a waggle of their abdomen (for a more detailed and nuanced discussion see Barron and Plath 2017). The direction of the straight component of the movement corresponds to the direction of the food resources found by the insect (in relation to the axis between the entrance to the hive and current sun position), and the duration of the waggle movement indicates the distance to the resource.

In this context, the biosemantic approach notes the dance as a representation (and a signal, in this case) shared by the producer (the dancing bee) and the consumer (in this case in plural: multiple bees observing the dance). Features of the representation correspond to the external states of affairs (direction and distance towards the resource) which, when perceived by the consumers in normal conditions enables them to perform their proper function (specifies a suitable place for foraging for nectar, enabling the other bees to minimize time and effort spent for looking for proper food sources). What this means is that the representational *status* of a particular behavioural instance depends on how that behaviour is used by some other system (or more precisely, how that behaviour evolved to be used, as we will see shortly in the discussion of the notion of "proper function").

On the other hand, the representational *content* is fixed by the mapping or correspondence between the states of the behaviour and some other states, and the consumer's proper function this correspondence normally serves (e.g., this is how we can establish that the bee dance refers to the "location of food source" rather than the "location of a beautiful linden tree in full blossom", even if both of those objects coincide spatiotemporally—solving the proximal/distal content difficulties of Dretske's account). Fixing the function for analysis enables us also to distinguish between the directive and descriptive types of content, emerging out of the original pushmi-pullyu form.

In Millikan's view, directive representations have a proper function to guide a consumer to produce their own satisfaction condition (in a way that a blueprint specifies what needs to be constructed, see Millikan 1995). Descriptive representations, on the other hand, do not have a proper function of their own, but their satisfaction conditions emerge as a result of the *consumer's proper function*: more precisely, the content is true if (and only if) it correctly "adapts its interpreters [...] in the service of *their* proper function [as in the case of the bee dance—which indicates the location of the food source on this analysis, and not a command for other bees; Millikan (1995), p. 189]. Two central concepts in this account require further discussion: the notion of a"proper function" and of "normal conditions", which are both important elements of Millikan's definition.

The notion of "proper" functions has been introduced by Millikan in her discussion of the naturalization of the notion of biological function, in response to the limitations of previous proposals. Philosophers of science have approached the task of accounting for the broad use of functional notions in explanatory practices of sciences without reliance on the "suspect" teleological notions, such as "design" (cf. Dennett 1996), by employing the idea of "causal role" (I'm following here the account in Couch, n.d.).

This tradition goes back to the work of Ernest Nagel (1961), who initially proposed a way of reducing functional talk to the (then) standard deductive-nomological (D-N) view of scientific explanation (see Hempel and Oppenheim 1948). His account suggests that references to functional terms are "explanation sketches" that attempt to explain why a particular component—whose function is the target of explanation—is present in a system. A complete version of the sketch would then proceed to list the necessary conditions for the occurrence of the component (or activity) in the system (Nagel 1961, p. 405): in organisms, this means in particular the contribution of those systems to survival and reproduction.

This account has faced numerous criticisms: in particular, the redescription of function it offers does not track well with actual scientific practice (see Cummins 1975), leading to cases—e.g., when a particular trait becomes maladaptive due to the presence of a new predator—where the D-N account does not prescribe any function, while in scientific practice, the function would still be identified. Mark Couch suggests as an example the presence of wings in some species of insect, which stopped contributing to its survival due to a new airborne predator: Nagel's account would not classify them as having a function, while researchers might still—quite sensibly—claim that the wings contribute to the ability of insect to fly (Couch, n.d.).

In response to this and other limitations of Nagel's account (and the broader move away from the D-N view of explanation), Robert Cummins has advanced an alternative approach (see Cummins 1975, 1977, 1983). He approaches the functional explanations by identifying two stages: first, researchers perform a functional analysis, which decomposes the function into a set of simpler, "less problematic" (Cummins 1977, p. 272) dispositions (i.e., functions), such that their "organized manifestations [...] amount to the manifestation of the analyzed disposition" (Cummins 1977, p. 272). Second, researchers identify a physical structure capable of instantiating these constituent dispositions, to show that their functions are actually performed. The analysis can be further reiterated, and constituent functions at some level can be the targets of analysis at a different one. In practice, this process follows the research interests of the scientists involved. Cummins' function, then, identifies in what way the analysed component is causally contributing to some capacity [i.e., a set of behaviours or activities, such as flight, pumping blood, learning the maze; Cummins considers it a more primitive notion than "disposition", (1975), pp. 759-760] of the system.

While Cummins' focus on the causal contributions of components has shaped the subsequent work on biological functions, his account has faced several important criticisms. Millikan, in particular, has argued that this approach is too liberal in ascribing functions and fails to account for malfunction, which is commonly identified in a variety of sciences (e.g., in medicine or psychiatry this is of critical importance). The first point Millikan raises (1989a) is that biological systems commonly do many things concurrently: hearts, a favourite example in these debates, have a function of pumping blood, but the contracting and relaxing capacities of heart muscle cells which underlie this function at the same time cause a thumping noise. Now, we commonly do not qualify the thumping-producing capacity of hearts as their function, while on Cummins account it is impossible to distinguish this capacity from that of pumping blood.

Second point of criticism identified by Millikan (1989c), and further raised by Karen Neander (1991) refers to instances of malfunction which result e.g., from lacking or malformed components. If a child is born with a congenital heart valve disease, their heart may fail to pump blood effectively and thus not work. Under Cummins' account of function, the lack of valves, critical for the performance of the blood-pumping function in a healthy heart, would mean that the functional analysis fails, and the malformed heart should not have a function ascribed. But this does not agree with the use: we normally would still consider the heart as having the function to pump blood, even if for some reason it fails to execute it.

In response to these objections, several authors, most notably Millikan (1984; 1989a; 1989c), Neander (e.g., 1991) and Paul Griffiths (e.g., 1993) have defended the notion of "proper functions" and a broader etiological account of teleological notions. Importantly, the strategy of these authors differs from the ones discussed above in that they're not trying to account for the practice of functional explanation and the use of "function", but rather offer a primarily theoretical account of what function is (Millikan 1989c).

The etiological account does not shy away from relating function to design. However, its proponents view design in biological systems in terms of their evolutionary history, and fitness benefits a particular component or activity bestows on the system that have led to that component or activity being selected for and reproduced. As such, an item A has a proper function F if (1) A originated as a "reproduction"[3] from some prior item due (at least in part) to the properties of that previous item performing F in the past, and this performance causally led to the origination of A, or if (2) A originated as a product of some historically prior item that had F as a proper function, and under normal circumstances, the item has performed F by means of producing an item like A (see Millikan 1989c, p. 288). Items falling under the first condition are considered as having direct proper functions, while items in the second category have derived proper functions.

Consider again the example of the human heart: it would be considered as having a derived proper function of pumping blood, as it is not directly copied, but produced as part of the larger system. Human body has the direct proper function of pumping blood, and normally it executes the function by producing hearts. This production enables successful execution of the function and the

---

[3] This is a technical term that Millikan defines in great detail in Millikan (1984). For the purpose of the current discussion, this can be understood as copying, or copying a copy, including the cases of biological reproduction were genetic material is being copied.

reproduction of the whole organism is possible (in part) because of that success. At the same time, the production of thumping noise is not a proper function neither of the body nor of the heart, as it is likely causally irrelevant to the reproduction of the organism. Further, if the organism is not developing in normal conditions (e.g., due to a harmful genetic mutation), it can fail to produce a heart that is able to pump blood effectively. However, the malformed heart still has the derived proper function of pumping blood, as if it were to develop in normal conditions, it would pump blood.

The final element that requires discussion is the idea of "normal conditions", that forms a part of both the teleosemantic account of content and of the notion of proper function. She contrasts her understanding (capitalized as "Normal") of this concept with the prevalent tendency for "handwaving, when things get rough, toward the relevant *ceteris paribus* clauses under the heading 'normal conditions' " (Millikan 1989c, p. 300). In Millikan's account, Normal conditions refer to the conditions in which organisms under consideration have *historically* found themselves, and in particular: conditions that have historically occurred, and when the dispositions under consideration have been adaptive. What this means is that "Normal" in no way refers to any concept of frequent or average conditions. Indeed, a particular behaviour might have a proper function even if it has seldom occurred in the (evolutionary) past—what matters is that in those instances it did act adaptively, and the contemporary instance of that behaviour is preserved because of that past occurrence.

In the context of communication, Millikan considers the fact that some types of animal signalling that have their proper function of alarming conspecifics against a predator, under most occasions constitute what we would call false alarms: i.e., the signals are triggered in the absence of any danger. The etiological account still ascribes them the function of warning against predators, as historically these instances of the behaviour have acted adaptively (when the predator actually was present) and the contemporary manifestations are caused by that adaptive character (alarms allowed the animals to escape, which led to their preservation and enabled eventual reproduction).

While Millikan's approach to function has faced some criticisms and seen some further developments, it has been thoroughly defended and in result, the etiological approach has dominated the debates regarding the naturalization of function in biology. While this perspective has not been explicitly applied to the analysis of instances of biological signalling outside of animal communication (including language), the close intertwinement of function and meaning emerging from this work, and the category of pushmi-pullyu representations discussed above, suggest that it might help in understanding how "meaning" and function are identified by researchers studying cellular signalling. In this way, the teleosemantic approach offers important benefits for the current project, compared to some other prevalent approaches that are primarily concerned with linguistic reference.

For example, the prominent mental files approach (see Récanati 2013), which views reference as direct, in virtue of the contents of the relevant mental file (in the sense of a file in a filing cabinet), requires the presence of epistemic properties (the "epistemically rewarding relation") that we might not be able to usefully ascribe to as simple biological systems as cells (even if this approach has some benefits in terms of how it combines semantics and pragmatics and its non-descriptivist focus). Similarly, Kit Fine's coordination theory (Fine 2007, 2017b, 2017a) views a meaning of a particular (e.g., an individual representation) at least partly resulting from its coordinative role with regard to other meaningful expression. As a form of semantic holism, it focuses on higher level problems of semantics that do not translate onto the kinds of content that are of interest here (though I'll come back to the holistic point below, in subsection 5.5.1). For this reason, the teleosemantic approach seems to be uniquely suited for the following analysis.

5.2 NATURAL(IZED) ACCOUNTS OF CONTENT

In order to uncover the conceptual network related to semantics across the patches of BIOLOGICAL COMMUNICATION, I will query the corpus using the naturalized accounts of content and meaning which most significantly figure in contemporary debates in philosophy of mind. They all largely follow the general tradition of Ruth Millikan's biosemantics (e.g., Millikan 1984; Millikan 2004; see Schulte and Neander 2022). I have introduced above the Millikan's version of the teleosemantic approach, which remains the main frame of reference for the majority of the discussions surrounding meaning in philosophy of mind. Indeed, in what follows, I review six recent proposals that in various ways build on the teleosemantic tradition—three more directly embedded in that approach, and three accounts exploring different extensions of the teleosemantic framework. The selection was meant to give a thorough overview of natural accounts of semantics that are applicable to biological communication. Each review is geared towards distilling a succinct (if at times simplistic) summary of the view that can be used for querying the corpus.

5.2.1 *Teleological accounts*

Causal-informational theory of content

Karen Neander (Neander 2012, 2017) builds on Millikan's theory, but rejects Millikan's focus on the consumer as responsible for establishing the content.[4] Instead, Neander (2012, p. 22) pursues a "hybrid idea [...] that representation-producing systems have the function of producing states that carry natural information. Then the 'aboutness' of content could be explained as originating with the 'aboutness' of information and the 'normativity' of content could be explained as originating with the 'normativity' of functions."

To this end, she introduces the notion of response function: she postulates that some biological systems—as she discusses the problem of content in the context of mental representations, she focuses on sensory systems—have a proper function "to respond to something by doing something" (Neander 2012, p. 23). This move enables Neander to claim that inputs and causes can impact the meaning. The view preserves the standard teleosemantic claim that the effects of biological mechanisms have direct impact on fitness (and thus are uniquely ascribed functions). However, considering the response functions of various biological systems allows Neander to enhance the standard view with the sensitivity to

---

[4] Importantly, Millikan's approach is slightly less consumer-focused as it tends to be interpreted. For instance, Millikan (1984), p. 98 notes: "An intentional icon Normally stands midway between two cooperating devices, a producer device and an interpreting device, which are designed or standardized to fit one another, the presence and cooperation of each being a Normal condition for the proper performance of the other." Hence, the meaning is dependent on the functions of both producer and consumer. I'm thankful for Ruth Millikan's thoughtful comments on the interpretation of her work in this chapter.

input causes, which translates into differentiated effects the response functions produce. In this way, her view is able to account for how inputs can influence the fitness.

With the notion of response function at hand, Neander is able to explain the ascriptions of information content, which Millikan rejected. To this end, she considers specifically the function of producing states that carry natural information. While she agrees with Millikan's criticisms of Dretske's account of informational teleosemantics and rejects the understanding of information in terms of correlation, she advances a distinct account of information in terms of causation (at the token level): we say that a particular response, R, carries natural information about event C, if C is a token-level cause of R (Neander variously calls this "singular causal indicative" or simply "natural" information, Neander 2012, p. 27). This notion of information escapes the formal apparatus of information theory (i.e., information transmission, as discussed in chapter 4), which is mathematically bound to consider only type-level properties. Nevertheless, Neander's notion of natural information does allow for accounting for several problems regarding the content of representation, e.g., rejecting ascriptions of content to spuriously correlated events and allowing for ascriptions of information that is not used or does not contribute to fitness.

This is a significant difference compared to standard teleosemantics, as it allows for ascribing the content to the producer. However, this is in fact a hybrid approach, as both the normativity and fixing of content relies on a standard teleosemantic story: as visible in her discussion of distal and proximal content (Neander 2017; see also Schulte and Neander 2022), Neander relies on the etiological considerations to argue that the content of a mental representation (or an utterance, for that matter) is a distal property rather than some more proximal element of the causal chain, precisely because the response function was selected for the fitness benefits it bestows in response to that distal content (in Normal conditions, in the sense discussed above), rather than to the proximal causes, i.e., its proper function. E.g., the vervet monkey alarm calls mean the presence of the predator/the need to escape because in normal conditions they are a response function to the presence of the predator (a distal cause), rather than the producer's belief that a predator is present (a more proximal cause)—even if in numerous cases they are caused by a mistaken belief of the producer, with no predator present and no need to escape.

Varitel semantics

Nicholas Shea's (2018) proposal develops the standard Millikan approach in several ways. Similarly to Neander, he strives to include informational notions in a teleosemantic account, but, different from her, Shea accepts a more standard, correlational notion of information. The term "varitel" semantics is meant to highlight the inherent pluralism of his approach: while offering a general account of content, he notes that two components are variable, leading to multiple "non-overlapping varieties" of content (Shea 2018, p. 43). These variable components are exploitable relations (within the representation/signal, or, more precisely, between the elements of the mechanism producing the representation) and (teleosemantically-understood) functions involved. Exploitability in this context means that a particular relation between the internal (representation or signal) and external (referent) states can be accessed and used as a representation.

Overall, he notes that representational systems ultimately produce "outcomes" (i.e., behaviour; Shea calls them also "task functions", developing his own version of the account of proper functions for this purpose) that are robust (with respect to variation in the inputs and in the conditions), stabilized (by a process of natural selection or individual learning) and that rely on a mechanism with "internal components which stand in for exploitable relations to aspects of the environment that are relevant to achieving an outcome" (Shea 2018, p. 51). While robustness and stability allow us to account for normativity, "aboutness" of representations is specified by the representing system's internal organization.

The internal organization that Shea postulates—and defends by looking at cases from cognitive science—is construed within the mechanistic theory of explanation (Machamer, Darden, and Craver 2000). He considers the problem of how a particular representational system is able to produce contentful representations, and proceeds by identifying the internal organization of the components of the mechanism producing those representations and studying their relations. In his view, as already mentioned, contents of representations rely on exploitable relations between the components of the mechanism. He lists two possible types of those relations (though this is not meant as an exhaustive list): *correlational information* and *structural correspondence*. Hence, the representation may rely on the correlational information carried by the components about features of the world—e.g., in "simple" cases of animal communication (see Shea 2007a). Alternatively, what matters is that the internal components are organized in such a way that there is structural correspondence between that organization and the organization of the representation's target (as we will see below, this relationship that can be formally described with Barwise and Seligman's notion of an infomorphism, introduced in subsection 4.1.3). The standard example in philosophy of cognitive science is that of cognitive maps (see Tolman (1948), O'Keefe et al. (1998); E. I. Moser, Kropff, and M.-B. Moser (2008); Bechtel (2016); Gładziejewski and Miłkowski (2017)). In this way, Shea introduces into the teleosemantic account a dependence on the input, which allows for more standard ascriptions of content than a strictly consumer-based theory would demand. At the same time, he advances significantly beyond the previous proposal by explicitly allowing for a plurality of (types of) content.

Non-etiological accounts

Alternative developments of the teleosemantic approach drop the etiological account of function that has been a feature of all the views discussed so far.

Most notably, this is the step that Gualtiero Piccinini takes. He advances a view of representation, initially focusing on sensory representations, which relies on the goal-contribution account of biological function (Piccinini 2021, ch. 3). According to this account, functions are regular contributions to biological goals, such as survival or reproduction. While still partially historically determined, this account of function seeks to lessen its dependence on historical facts (and the associated difficulties, Schulte and Neander (2022)). In result, Piccinini offers the following definition of representation for sensory systems:

> "A state (or signal) S within an agent's representational system R *indicatively represents* that C [if and only if:] A function of R is to produce S, so that S carries natural semantic information that C, S is one among a range of similar states that map onto a range of similar external states, and S can guide the agent's behavior with respect to the fact that C." (Piccinini (2021, p. 263), notation changed for coherence)

His view, similarly to Shea's, accepts a more standard view of information. Specifically, by "natural semantic information" Piccinini builds on Grice's notion of natural meaning (discussed in subsection 4.2.4), connecting it with the standard mutual information. Natural semantic information is carried by a particular signal if that signal carries mutual information about a source, and the occurrence of a state of the signal raises the probability of a particular state of the source (Piccinini 2021, p. 130). Crucially, his understanding of natural information is restricted to this correlational meaning: the signal or representation carrying natural information can serve only the purpose of *tracking* its source (see Piccinini 2020). The further conditions he specifies require that the signals form a larger, organized set or be part of a mapping (preventing "accidental" ascriptions of representation), and that they be actually consumable for guiding behaviour, a condition stemming from the definition of function he accepts.

Piccinini offers also a slightly different account of motor representations—what he calls *imperative representation* (analogously to Millikan's distinction of descriptive and directive represenations, see above). While Piccinini further extends his account to nonnatural meaning (via the discussion of offline mental simulation, see Piccinini 2020), he considers biological communication cases (he lists vervet monkey calls) as instantiations of arbitrary signalling systems, which in his view are restricted to natural meaning: "[a]s far as I know, the function of [vervet monkey's] alarm calls is to track and communicate the presence of predators. If this is their only function, then [they] have only natural meaning" (Piccinini 2020, p. 280). Beyond this brief remark, however, Piccinini does not consider such cases, focusing on explaining the semantics of mental representations.

In a similar vein, Bence Nanay (2014) proposes a theory which relies on his modal view of function (Nanay 2010). He restricts his account to "pragmatic representations", i.e., those which are most directly involved in guiding action (see Nanay 2014). In his view, a representation R represents that C if and only if R has the function of carrying information that C, where function is ascribed to a trait or behaviour, if that trait or behaviour would contribute to the organism's fitness (which is the modal definition, see Schulte and Neander 2022). Nanay's restriction of the scope of his account to pragmatic representations, as he understands them, implicitly excludes from consideration the majority of biological signals, which cannot be duly considered "immediate *mental* antecedents" of action (Nanay 2014, p. 805; emphasis added). In result, his account offers little help in the context of the current project.

5.2.2 *Beyond teleosemantics*

Recently, several accounts have been advanced which extend to varying degrees the teleosemantic approach and more directly rely on the formal apparatus of information theory. While Shannon famously stressed that "semantic aspects of communication are irrelevant to the engineering problem" (Shannon and Weaver 1964, p. 31), the importance that the formalism he advanced has attained has led many researchers to search for a theory of semantics that is in close agreement with the information theory's conceptual and mathematical background. Indeed, already Warren Weaver noted that Shannon's statement "does not mean that the engineering aspects are necessarily irrelevant to the semantic aspects" (Shannon and Weaver 1964, p. 8). Here, I look at three recent proposals in this vein, reviewing them in chronological order. Importantly, as will become clear, they aim to answer a distinct set of problems, and while they rely on the teleosemantic approach to varying degrees, they remain compatible with it (Miłkowski's correspondence theory is perhaps the closest to standard teleosemantics). Nevertheless, the conceptualization of meaning that emerges from this, more information-oriented perspective provides an important set of analytical tools for studying the use of semantic concepts in scientific literature, precisely because of the close relationship to the standard information theory.

Causal necessity account

Artemy Kolchinsky and David Wolpert (2018) take information theory as their point of departure. They develop their view by coupling the theory to the framework of non-equilibrium statistical physics. Their goal is to provide an account of semantic information suitable for the study and development of *autonomous agents* (see Harrison, Rorot, and Laukaityte 2022): an explicitly ahistorical view, "applicable to any physical system coupled to an external environment, whether a rock, a hurricane, a bacterium, or a sample from an alien planet" (Kolchinsky and Wolpert 2018, p. 2). Importantly, this broad scope is justified by their attempts at quantifying the amount of semantic information. This eventually leads them to the (quite understandable) conclusion that rocks or hurricanes have low or very low amount of semantic information available at their disposal (Kolchinsky and

Wolpert 2018, p. 3)—an important result in establishing the adequacy of their proposed account.

Briefly, the causal necessity account of semantic information states that semantic information is “the information that a physical system has about its environment that is causally necessary for the system to maintain its own existence over time” (Kolchinsky and Wolpert 2018, p. 1). There are several aspects of this definition that require unpacking (the original article does so by providing a useful formalism, however in the current context I will focus on providing the interpretation and motivation for the convention they adopt):

1. First, the notion of information at play refers to purely syntactic information that can be further specified as needed (see below).

2. Second, the definition applies to physical systems which maintain their existence over time (given a particular timescale). While this is a broad class (as visible from the inclusion of rocks and bacteria above), this second condition immediately points out non-equilibrium steady state systems, and in particular recursively self-maintenant systems (see Bickhard 2009; cf. Harrison 2023; McGivern 2020; Seth 2025), as a class of particular interest.

   This broad scope is critical for how we understand “maintaining its own existence”: the authors operationalize this notion by introducing a class of “viability functions” which are intended to measure the “degree of existence”. While a particular viability function can be chosen according to a research context (see below), Kolchinsky and Wolpert decide on defining it in terms of Shannon entropy of a distribution of states which—within the thermodynamic approach to life they adopt (going back to Schrödinger 1944)—they define as providing a measure of probability that the system occupies one of the thermodynamically “viable” states (i.e., a state in which energetic conditions are such that the system remains away from thermodynamic equilibrium). Hence, “maintaining existence” amounts to a non-negative change in the value of the viability function for the system.

3. Finally, perhaps most importantly, the notion of “causal necessity” is here explained in broadly counterfactual, modal terms, which are grounded in the more standard considerations of causal analysis (in the sense of Pearl 2022, see also the discussion in subsection 3.5.1).

   Kolchinsky and Wolpert consider possible interventions on the system which scramble (i.e., arbitrarily rearrange) the relationship between the system and its environment, reducing the amount of syntactic information in the system. The value of information can be specified as the difference between the actual and scrambled viability of the system. In this context, we can look for a particular optimal intervention, which would be the maximal scrambling that preserves the viability. This value givess as the least amount of (syntactic) information necessary for the actual viability of the system.

> This notion of optimal intervention is crucial to the semantic content of a particular state, which is defined "as the conditional distribution (under the optimal intervention) of the environment's states, given that the system is in state x" (Kolchinsky and Wolpert 2018, p. 4).
>
> In less technical words, the semantic content of a particular state—e.g., a signal—is the set of correlations causally necessary for maintaining the viability of the system. If any of those correlations would be misleading or missing from the signal (as a result of the scrambling), the state would not contribute as much to the system's viability.

Consider again the case of vervet monkeys and the semantic contents of their calls (a standard description of this case is provided by Seyfarth, Cheney, and Marler 1980). A particular signal, say an `eagle-call`, has some correlational relationship to the environment state, i.e., it will correlate with situations when there's an eagle hunting in the sky. If we scramble its relation to those environmental properties (i.e., consider a signal in an environmental state that it is not normally related to), we can consider how much information it still has and how the scrambling impacts the viability function. If we consider the call in the `no-predator` environment state, it will have no informational content, and some negative impact on intrinsic viability, as it will likely cause an escape response, increasing the energy expenditure. Now if we consider the call in the `leopard` state, the `eagle-call` might have some informational content (i.e., informing about the presence of the predator—although it is unclear whether this would be an adequate analysis of the signal, see the discussion of coarse-graining below in subsection 5.5.2), but will have a largely negative impact on the monkeys' viability, causing them to search hiding in the bushes, precisely where the leopard is lurking. In this case an optimal intervention would be one that swaps one eagle to another: the call doesn't differentiate whether the predator is `eagle`$_1$, which hunts our vervet monkey troop on even days of the month, or `eagle`$_2$ which hunts on odd days. In states `eagle`$_1$ and `eagle`$_2$ the call has the same impact on viability, which motivates our calling it the `eagle-call` in the first place. At the same time, if we consider a viability function over a larger timescale and the whole troop, we might arrive at a slightly different consideration: for example we can imagine that the `eagle-call` in the `no-predator` environment state, instead of just impacting negatively energy expenditure, has the benefit of allowing the young to practice their understanding of calls and finding a proper escape route. Under such circumstances, we would perhaps consider the signal as a `run-to-the-bushes-call`. Now, the causal necessity account explicitly allows this kind of variability, biting the bullet on the indeterminacy of content, at the same time assuming that there is some "true" meaning of the signal for the monkeys: an intrinsic semantics, which we may attempt to approximate.

More broadly, this account aims at semantic information intrinsic to the system, yet the measures it offers are perspectival, as they hinge on several choices made by the researcher that is studying the system in question: the definition is provided so that various measures of syntactic information, of the viability function, and of conceptualizing the intervention, as well as the particular distinction of system and environment used or the scale investigated, are possible (see Kolchinsky and Wolpert 2018, p. 4). Hence, if we make different choices of measures of syntactic information or the viability function, we may arrive at widely different numerical values of semantic information within this formalism. Nevertheless, assuming that there is something like a viability function and syntactic information *for the organism*, the causal necessity view is able to account for the intrinsic meaning as well, even if we may not be able to access it epistemically.

S-vector semantics

Alistair Isaac (2019) notes the historical relations between the mathematical apparatus of information theory as developed by Shannon, and the almost concurrent achievement of an analogous formalism by Turing and his colleagues (see the analysis in Good 1979). The first one explicitly addressed only the "engineering problem" of transferring a signal over a noisy channel. However, the other has addressed precisely the problem of the content of a given string, as part of his work on decoding the Enigma during the Second World War. Based on this historical parallel, Isaac proceeds to claim that there are semantics latent in Shannon information. He develops a theory, called "s-vector semantics", grounded in earlier work of Brian Skyrms (2010) and closely related to the proposal of Andrea Scarantino (2015). His proposal in important respects parallels the pragmatic considerations motivating the distributional approach to meaning prevalent in computational linguistics, that have been discussed in chapter 2. While these two approaches share an underlying formalism, Isaac's theory is particularly geared towards the issues prevalent in philosophical debates on (natural) meaning.

Isaac considers the natural meaning of a signal in terms of a change of the subjective probability of a target event (property, state of the world) of the recipient of the signal, conditioned on that signal (I will omit here some formal details of the account, to focus on the motivations and interpretation, as in the previous section). To provide an example, a troop of vervet monkey foraging the trees has a certain prior probability of encountering a predator, either a leopard hunting on the ground, a snake, or a bird of prey in the sky. However, once a sender spots a leopard lurking in the bushes and gives out the `leopard-call`, a recipient monkey can raise its subjective probability associated with there being a land predator in the vicinity, and act accordingly [e.g., run up the tree; the standard description of this situation is provided by Seyfarth, Cheney, and Marler (1980)]. On Isaac's proposal, the warning call has a natural meaning of there being a land predator precisely because of the particular log probability ratios of different events it conveys to the recipient. Crucially, this goes beyond the standard Shannonian quantification that (assuming a uniform distribution of the probabilities of different type of predators) would ascribe the same *amount* of information to distinct calls. S-vector semantics, on the other hand, is able to distinguish the content of the calls being either type of predator.

The notion of s-vector is a generalization of the concept illustrated, namely of log probability ratios between events. Under some standard formal assumptions (e.g., a finite state space, adequacy of an ergodic Markov process as the model of the events—see the discussion in subsection 4.1.2 and 4.2.3), Isaac starts with the definition of the amount of information in a single event, e, given a probability measure P. He proposes to quantify it as:

$$I(e) = -\log P(e)$$

as the function captures several intuitive properties of information: information is always positive, it is additive, and the more certain an even is, the less information it provides (and vice versa).

He uses this function to describe the change in information about an event C provided by some other event, R (i.e., the amount of "evidence" R provides about C, building on Turing's work, see Good (1979); Isaac (2019), p. 110) as follows: consider a certain finite set of events C and an algebra $\mathcal{C}$ over that set such that if $c, r \in \mathcal{C}$, then $c, r \subseteq C$. We have then:

$$I(c : r) = I(c) - I(c|r) = -\log P(c) + \log P(c|r) = \log \frac{P(c|r)}{P(c)}$$

Now we can generalize this by considering a vector $\nu(r)$ for all elements $c_{1\ldots n} \in \mathcal{C}$ such that for any $c_i \neq r$:[5]

$$\nu(r) = \langle \log \frac{P(c_1|r)}{P(c_1)}, \log \frac{P(c_2|r)}{P(c_2)}, \ldots, \log \frac{P(c_n|r)}{P(c_n)} \rangle$$

Isaac calls this "the basic semantic object implicit in standard information theory" (Isaac 2019, p. 111). This is the definition of an s-vector for the signal (or representation) r, and it encapsulates the total information content of that signal. The theory of s-vector semantics Isaac outlines states that r (naturally) means $\nu(r)$.

Consider the vervet monkey calls in a simplified fashion. We have a set of four events: $C = \{$`no-predator`, `leopard`, `snake`, `eagle`$\}$. Assuming that events are mutually exclusive, the algebra $\mathcal{C}$ will also form a four element set (of four single-element sets for each element of C). Further, there is a probability measure P which assigns some non-zero probabilities to the elements of $\mathcal{C}$.

According to this formalism, a meaning of a certain call, e.g., `eagle-call`, is a vector of log probability ratios:

---

[5] While Isaac's approach allows for any event to be ascribed semantic content (i.e., any event can be considered as a signal or a representation) a clear distinction between signals (r) and events (c) they refer to will aid in this short exposition—hence, it what follows I will assume that these are distinct classes of events, but this not an element of Isaac's view (cf. Isaac 2019, pp. 113-114).

$$\begin{aligned} v(\texttt{eagle-call}) = \langle\, & \log \frac{\texttt{P(no-predator|eagle-call)}}{\texttt{P(no-predator)}}, \\ & \log \frac{\texttt{P(leopard|eagle-call)}}{\texttt{P(leopard)}}, \\ & \log \frac{\texttt{P(snake|eagle-call)}}{\texttt{P(snake)}}, \\ & \log \frac{\texttt{P(eagle|eagle-call)}}{\texttt{P(eagle)}} \rangle \end{aligned}$$

We call this an "eagle call" because all log probability ratios in this case would be negative, except for the value of $\log \frac{\texttt{P(eagle|eagle-call)}}{\texttt{P(eagle)}}$: the call decreases the probability of all events except for there being an eagle attempting to prey on the troop (which follows from the assumption of the events being exclusive).[6] A `leopard-call` would have a different meaning, precisely because a vector $v(\texttt{leopard-call})$ would consist of negative elements, except for the value of $\log \frac{\texttt{P(leopard|leopard-call)}}{\texttt{P(leopard)}}$.

While tightly connected to the amount of information carried by the signal, this formalism allows for distinguishing the contents of two different messages which carry the same amount of information, as would be the case if all predators were equiprobable. Crucially, this is not a matter of pure correlation: the meaning of a signal derives from the changes of probabilities of events enacted by the signal upon the receiver, not from any consumer-independent value of mutual information, as in Shea's account discussed above.

Isaac's account focuses on consumer-meaning, but it omits the discussion of *how* the recipient uncovers the meaning of a particular signal. While discussing the instrumental distinction of events into messages and (potential) referents (that I have also taken advantage of above), he notes only that his perspective is "egalitarian" about information content and "[i]t is only when organisms use the information available in nature, by detecting some events with their perceptual organs, and responding to other events in ways that reflect their importance for survival, that it makes sense to model correlated events as sorted into signs and signifieds" (Isaac 2019, p. 114). His mention of survival, and his earlier insistence of compatibility of his account with the teleosemantic perspective suggests that this "sorting", and more broadly, interpretation of signals, relies on some sort of etiological teleosemantic account. Otherwise, this account doesn't put any clear distinction between the meaning of signals and natural signs (cf. subsection 4.2.4).

While Isaac shows how this account can be applied to instances of linguistic meaning (and a host of philosophical problems, including the problem of error), the discussion so far should suffice in showing its relevance for biological communication more broadly. While we might lack epistemic access to individual cells' partitioning of the world into states, we can apply an s-vector semantics also

---

[6] In fact, the call structure of vervet monkeys is more complicated than this toy description allows, see (Price et al. 2015; Dubreuil et al. 2025)

to cell-cell communication since we are able to note how the presence of signals modulates probabilities of particular futures (e.g., developmental trajectories) for individual cells and larger populations.

Correspondence theory

Marcin Miłkowski (2020; 2023b) attempts a similarly minimal (in terms of the employed assumptions and conceptual apparatus) account of semantic information, based on the concept of similarity. Understanding similarity in terms of correspondence, he offers an account inspired by Tarski's correspondence theory of truth. If we specify the meaning of a piece of information in terms of its satisfaction conditions (i.e., the conditions or states of affairs that must obtain to fulfill the piece of information), the main claim of Miłkowski's correspondence theory is that "a piece of information is true if and only if it corresponds to what it is about" (Miłkowski 2023b, p. 486).

To unpack this, we need to focus on the notion of "correspondence". While "similarity" is central to his account, Miłkowski considers it primarily in terms of structural correspondence between representation vehicles and representation targets (allowing for other forms of similarity, and a representational pluralism more broadly). Structure of representation vehicles corresponds, roughly, to the independent degrees of freedom of its physical implementation (cf. MacKay 1969). In characterizing the correspondence relation, Miłkowski adopts the perspective of Jon Barwise and Perry Seligman (1997, see the subsection 4.1.3 for a broader discussion of their formalism).

Miłkowski employs their definition of classification to define the relata of his concept of correspondence, and generalizes the notion of infomorphism, i.e., a pair of functions translating between classifications, into that of *infocorrespondence*, which lifts the surjectivity condition of the infomorphism (i.e., allows for some tokens not to be assigned types), allowing for different degrees of correspondence. This generalization is especially relevant if we consider classifications with fuzzy type membership. Importantly, this definition does not require accepting any single implementation of the pair of functions $\langle f^{\Sigma_R}, f^C \rangle$ specifying the infocorrespondence, allowing for representation to rely on various species of similarity, in a similar vein as we saw in Shea's varitel semantics above. At the same time, Miłkowski quite strictly views these infocorrespondences as produced by causal, mechanistic processes, which underscores that the formal description in terms of Barwise and Seligman's framework is intended to track actual interactions of the representational system.

Initially, for the purpose of generality (e.g., application to the instances of mental representations and representations investigated by philosophers of science), Miłkowski sets aside the question of how the target of representation is selected. However, in a recent paper with Wojciech Mamak (Mamak and Milkowski 2025), they further develop this aspect of the theory. They posit that correspondence-based information in cognitive systems is exploited by a class of representational mechanisms. These are computational mechanisms that have the (proper) function to play a representational role in the system they are part of. They posit additional conditions drawing on the previous work on computational mechanisms (Miłkowski 2013), and the account of structural representations proposed by Paweł Gładziejewski (2016; Gładziejewski and Miłkowski 2017). These allow them to introduce additional conditions on the representational mechanisms, in particular their ability to detect errors (see Bielecka and Miłkowski 2020, 2024). Overall, this extension of the account allows them to respond to worries regarding the selection of representation target and highlights the physical character of representation vehicles (as their analysis of the Representational Similarity Framework shows, see also subsection 1.3.2).

To summarize, what this work suggests is that relevant correspondences are picked out by the components of signals (representation vehicles; in the new mechanist sense, as outlined above in subsection 5.2.1) that are causally contributing to the success of the recipient's behavioural or inferential strategy, and—crucially—are contingent and prone to error which can then be detected (if the strategy fails or simply does not achieve *full* success) and corrected (so that future uses of the representation of signal lead to a greater success).

Let's return to the example of vervet monkey calls to see how this theory fares beyond its intended scope. Assuming again the simplified description which identifies four possible states of the environment, $C = \{\texttt{no-predator}, \texttt{leopard}, \texttt{snake}, \texttt{eagle}\}$, we have $\Sigma_C = \{\texttt{no-predator-present}, \texttt{predator-present}\}$ and a binary function $\models_C$ that assigns events in C their types from Σ_C (this should be self-explanatory in this example). Analogously, we have a set of calls $R = \{\texttt{no-call}, \texttt{leopard-call}, \texttt{snake-call}, \texttt{eagle-call}\}$, an appropriately defined set of types, $\Sigma_R = \{\texttt{no-call-produced}, \texttt{call-produced}\}$, and an obvious function $\models_R \colon R \rightarrow \Sigma_R$. Note that even without a formal proof, the infocorrespondence between $\mathbb{C}$ and $\mathbb{R}$ is quite clear, leading us to assume that calls can carry semantic information about so-defined states of the world, with the truthfulness or accuracy of individual instantiations being directly evaluable within the correspondence framework.

5.2.3 *Distilling definitions*

This discussion of different views on how semantics can be naturalized was necessary to present the contemporary landscape of debates on this topic in philosophy. The teleosemantic approach clearly dominates the discussion, offering apparently the most promising perspective on the problem, and even if some approaches set its assumptions aside, they often highlight compatibility (see subsection 5.2.2 for an example), an important rhetorical point aimed at garnering support, which clearly highlights the centrality of teleosemantics. In what follows, I will use those perspectives to identify passages in the corpus which are concerned with the semantics of communication across different levels.

For this reason, I have distilled down the definitions discussed extensively above into slightly simplistic, single-sentence versions which can be used for querying the corpus (presented in box 5.1). Notably, I try to avoid technical notions which may not be represented in the corpus. Also for this reason, all definitions refer to the meaning of a "token" which is located near words such as "signal", or "message" in the embedding model used, but is a more generic term and as such provides more diverse results (e.g., in preparatory queries, when the word "signal" was used, the results were heavily skewed towards examples that explicitly mentioned "signalling" in various ways).

Box 5.1: Definitions of meaning based on the discussed theoretical positions, used as queries for semantic search.

These queries were used in the second study, as mentioned in chapter 2. The study follows the same methodology laid out in subsection 2.5.2, as the study discussed in chapter 3, though the total number of results is smaller due to the smaller set of queries. Similarity score values are mean cosine similarities between the query and the 10,000 results from the corpus produced by that query, with standard deviations, rounded up to four decimal places.

1. ID: 0_Biosemantics (Millikan)
   QUERY: During communication, the meaning of a token is established by what, upon the token being received, enables the recipient to perform effectively a function contributing to survival, due to a mapping between the token and what it refers to.
   SIMILARITY SCORE: $\mu = 0.3572$, $\sigma = 0.0615$
2. ID: 1_Causal-informational (Neander)
   QUERY: During communication, the meaning of a token is established by what causes the sender to produce the token in accordance with a particular response pattern, informing the recipient about the cause of the token.
   SIMILARITY SCORE: $\mu = 0.3668$, $\sigma = 0.0570$
3. ID: 2_Varitel (Shea)
   QUERY: During communication, the meaning of a token is established by the structure of the signal and what it relates to, through a correlation or structural correspondence, and the goal-directed, stable, robust behavioural outcomes it produces and has produced in the past.
   SIMILARITY SCORE: $\mu = 0.3911$, $\sigma = 0.0488$
4. ID: 3_Non-etiological (Piccinini)
   QUERY: During communication, the meaning of a token is established by what it carries mutual information about, allowing the recipient in a systematic way to act appropriately, and improving the performance.
   SIMILARITY SCORE: $\mu = 0.3616$, $\sigma = 0.0775$
5. ID: 4_Causal necessity (Kolchinsky and Wolpert)
   QUERY: During communication, the meaning of a token is established by the information that is causally necessary for the recipient to maintain its own existence over time.
   SIMILARITY SCORE: $\mu = 0.3718$, $\sigma = 0.0601$
6. ID: 5_S-vector (Isaac)
   QUERY: During communication, the meaning of a token is established by how it changes

the perceived probabilities for the recipient.
Similarity Score: $\mu = 0.3778$, $\sigma = 0.0699$

7. ID: 6_Correspondence (Miłkowski)
   Query: During communication, the meaning of a token is established by the structural correspondence between the token and its target, in such a way that the correspondence is causally relevant to the success of a recipient's strategy, and can fail and be corrected.
   Similarity Score: $\mu = 0.3888$, $\sigma = 0.0613$

5.3 RESULTS

The queries listed above were used for semantic search of the S2ORC subcorpus, following the procedure described in subsection 2.5.2. Search resulted in 16,485 sentences and corresponding paragraphs, balanced across the queries and disciplines. As mentioned, these definitions lose a lot of nuance, and as such it is not directly possible to use this approach to analyse how well these accounts reflect the actual understanding of meanings in communication across the scientific fields (this will be partially done below, in section 5.6 through a close reading of results). However, they are sufficiently precise to pick out sentences that refer to semantics either explicitly or implicitly, as visible in the examples in box 5.2.

Box 5.2: Manually selected representative results for queries presented in box 5.1.

Manual selection was limited to top sentences across all areas and balanced across queries. Most of these sentences were produced by several queries, here they are classified with the best query (full list with scores can be found in the online supplement, available at: https://doi.org/10.17605/osf.io/ax4jm). The similarity scores are rounded up to four decimal places. The table has purely illustrative purpose, all the analyses reported in the chapter were conducted using the full data. The reference includes the CorpusID from S2ORC and the sentence number.

1. QUERY ID: 5_S-vector5
   SIMILARITY SCORE: 0.6020
   RESEARCH AREA: Psychology
   RESULT: This definition can be operationalized by building an instruction set that can generate the message.
   REFERENCE: 1750184, 500

2. QUERY ID: 6_Correspondence
   SIMILARITY SCORE: 0.5836
   RESEARCH AREA: Psychology
   RESULT: In this context a further organizational aspect deserving consideration concerns the possibility (see Agnati and Fuxe, 2000) that the communication pathway between a signal source and its targets could be under the control of a "modifier," responding to teaching signals that could originate, for instance, from environmental inputs.
   REFERENCE: 1724708, 145

3. QUERY ID: 5_S-vector
   SIMILARITY SCORE: 0.5811
   RESEARCH AREA: Psychology, Ecology and evolution
   RESULT: In humans and honey bees, the abilities of the listener directly constrain the meaning of information exchanged during an interaction.
   REFERENCE: 227241123, 108

4. QUERY ID: 5_S-vector
   SIMILARITY SCORE: 0.5508
   RESEARCH AREA: Neuroscience
   RESULT: In terms of communication, this then means that the awareness of other people's emotional states results in a kind of advantage for a human, in the form of a more

accurate generative model for the reasons behind other humans' actions.
Reference: 216044412, 62

5. Query ID: 5_S-vector
   Similarity score: 0.5473
   Research area: Molecular biology, Varia
   Result: We need to specify how association of meaning and signal in sending and receiving ends are done.
   Reference: 4880012, 58

6. Query ID: 6_Correspondence
   Similarity score: 0.5455
   Research area: Ecology and evolution
   Result: Transmitting cultural tradition (e.g., how to cope with unusually severe droughts) from one generation to the next requires a means of effective communication and therefore triggers evolution of speech.
   Reference: 4516380, 52

7. Query ID: 3_Non-etiological
   Similarity score: 0.5430
   Research area: Neuroscience
   Result: There are two subscales: "affirming communication", for example "we take the time to hear what each other has to say or feel", and "incendiary communication", for example "we walk away from conflicts without much satisfaction".
   Reference: 259157185, 151

8. Query ID: 5_S-vector
   Similarity score: 0.4997
   Research area: Neuroscience
   Result: This inter-communication is important in the development and maintenance of the nociceptive neurons' hyperexcitable state, i.e., central sensitization.
   Reference: 267157380, 228

9. Query ID: 6_Correspondence
   Similarity score: 0.4554
   Research area: Molecular biology, Neuroscience
   Result: Information is transmitted across compartments through dynamic protein interaction networks, which acquire, amplify, transmit, and eventually respond to signals that are channelled along the signaling cascades.
   Reference: 102348846, 16

10. Query ID: 2_Varitel
    Similarity score: 0.5293
    Research area: Animal behaviour, Ecology and evolution
    Result: As communication usually occurs in a network of several animals in signalling and receiving range of each other, the emotional state of a caller may influence the behaviour of several individuals, addressees in direct interactions and bystanders alike.
    Reference: 5566073, 24

11. Query ID: 5_S-vector
    Similarity score: 0.4931
    Research area: Animal behaviour, Ecology and evolution
    Result: Such signals can be understood as expressions of specific motivational states of the individuals, but there is no evidence that the subjects intend to alter the knowledge

state of the recipients.
Reference: 645745, 5

12. Query ID: 2_Varitel
    Similarity score: 0.4508
    Research area: Developmental biology, Microbiology
    Result: The review further discusses how these lysosome-derived signals in turn control autophagy, lysosome adaptation, cell proliferation and survival and immune cell function and many other outcomes, and how these observations support the view of lysosomes as a major signaling hub of the endocytic system.
    Reference: 204737631, 30

13. Query ID: 6_Correspondence
    Similarity score: 0.4475
    Research area: Developmental biology, Molecular biology
    Result: This signaling can lead to diverse series of functional and/or structural responses in the involved cells and tissues-the most extensively investigated of these are highlighted in each section of this Review (e.g. shape changes, migratory events and differentiation) (Fig. 1).
    Reference: 218985273, 126

14. Query ID: 5_S-vector
    Similarity score: 0.4420
    Research area: Developmental biology, Ecology and evolution, Molecular biology
    Result: These data indicate that although both buds have the potential to develop, signaling from a winner secondary bud creates a continuously repressive environment for the loser.
    Reference: 245217358, 178

15. Query ID: 2_Varitel
    Similarity score: 0.5030
    Research area: Microbiology, Molecular biology
    Result: In this approach, a signaling network makes some computations on the incoming signals and produces some responses at the outputs, accordingly.
    Reference: 202554259, 18

16. Query ID: 2_Varitel
    Similarity score: 0.4595
    Research area: Microbiology
    Result: Cell-cell communications, known as quorum sensing (QS) in bacteria, involve the signal molecules as chemical languages and the corresponding receptors as transcriptional regulators.
    Reference: 8196154, 0

17. Query ID: 5_S-vector
    Similarity score: 0.4911
    Research area: Molecular biology
    Result: The immune system and central nervous system have a vital means of communication through cytokines produced by immune cells.
    Reference: 234743417, 31

18. Query ID: 2_Varitel
    Similarity score: 0.4653
    Research area: Molecular biology
    Result: It is this information-processing capability of the signaling network that is

responsible for the context specificity of the cellular response.
Reference: 13280601, 13

19. Query ID: 2_Varitel
Similarity score: 0.4370
Research area: Ecology and evolution, Plant science
Result: This is the primary site for signal perception, the downstream trigger of the defense response and the cell fate decision under heavy metal stress.
Reference: 229324341, 108

20. Query ID: 2_Varitel
Similarity score: 0.4287
Research area: Molecular biology, Plant science
Result: Hence, positional signals communicated from neighboring cells largely control acquisition of cell identity.
Reference: 21548405, 17

5.3.1 *Topic modelling*

Paragraph model

To assess the results quantitatively, following the methodology adopted throughout the thesis, the paragraphs produced by semantic search were modelled with a BERTopic topic model. Similarly as in the other cases, a topic model of 107 topics (plus the outlier topic) was selected to achieve coherence-granularity tradeoff. The model was chosen through the modification of minimum cluster size (the hyperparameter that controls the HDBSCAN clustering algorithm in BERTopic—it was set to 25 (the smaller minimum cluster size, the more topics); no topic reduction methods were used this time as the automatic model size worked best). In this case, however, the resulting topics are not informative with regard to the contexts where semantics of communication are mentioned.

As visible in Table 5.1, the topic model relied primarily on the context of communication, similar to the one investigated in chapter 3. The topics are primarily distinguished by the different communicative processes that are mentioned across biology. An important conclusion that follows from this is that discussions of semantics of communication appear across the scales where communication is mentioned. There is however much less detail available in the model regarding *how* semantics are ascribed across the levels. One cue follows from the UMAP (dimensionality reduction) visualization of the paragraph embeddings (Figure 5.1), which indicates a clear distinction of two “clusters”: articles that refer to various forms of cellular signalling (either within larger organisms, or between unicellular life forms) are grouped together, and differentiated from paragraphs which discuss various forms of social and linguistic communication in multicellular animals (both in the context of human and non-human animals).

There is a small group of topics which are projected between the two larger groups (see Figure 5.2): topics 23 (top 3 keywords: ontology, terms, names), 25 (characters, species, margin), 96 (conservation, science, species), and 100

Table 5.1: 10 most frequent topics for the paragraph model (without the outlier topic) with base representations from the model. As the topics have not proven useful in analysis, the procedure for manual naming was skipped. The base representation is the default keyword selection algorithm based on c-TF-IDF (continuous term frequency—inverted document frequency, see chapter 2), limited to top 5 key terms.

Topic	Count	Base representation
0	590	cell, signaling, network, networks, cellular
2	455	conversation, speaker, speakers, alignment, turn
4	426	semantic, word, memory, words, processing
1	383	notch, cells, cell, signaling, differentiation
6	377	stress, drought, cold, abiotic, genes
8	345	message, messages, health, media, framing
3	341	integrin, actin, mechanical, cell, integrins
7	335	cancer, phosphorylation, activation, kinase, signaling
12	306	team, communication, employees, organizational, project
5	297	patient, patients, communication, training, clinical

(shape, geometric, topology) refer primarily to "meta contexts" (defining terms for ontologies and databases in topic 23, defining characteristics which indicate a species in topic 25, references to communicating the need for conservation in topic 96, or operationalizations for mathematical operations on biological data in topic 100). Topics 28 (probability, state, equation) and 31 (information, entropy, mutual information), on the other hand, indicate contexts which heavily rely on conceptual notions from various forms of probabilistic modelling of communication (topic 28), including the use of information theory (topic 31).[7]

While this data is inconclusive (especially considering the operations involved in the UMAP dimensionality reduction, which produces the visualizations), it suggests that there is some critical difference between how meaning is attributed at the level of individual cells, and at the level of multicellular organisms. The topic model also succeeds in picking out several "artifactual" contexts—for example the topic 101 groups together the "author contribution" sections from the corpus, picked out by semantic search. In this context one could also view the topic 34 (family, life, care) which is concerned with end-of-life care, and refers to communication as "meaningful" in the sense of significant, or impactful,

---

[7] This structure is not merely an artifact, as further confirmed by exploring different hyperparameters of the topic model: if minimal cluster size for the HDBSCAN clustering algorithm was set to values of 50 or 150, the resulting model consisted of, respectively, 2 or 3 topics (the third topic is significantly smaller), which showed an analogous pattern. The visualizations generated by these models are included in the online supplement, available at: https://doi.org/10.17605/osf.io/ax4jm.

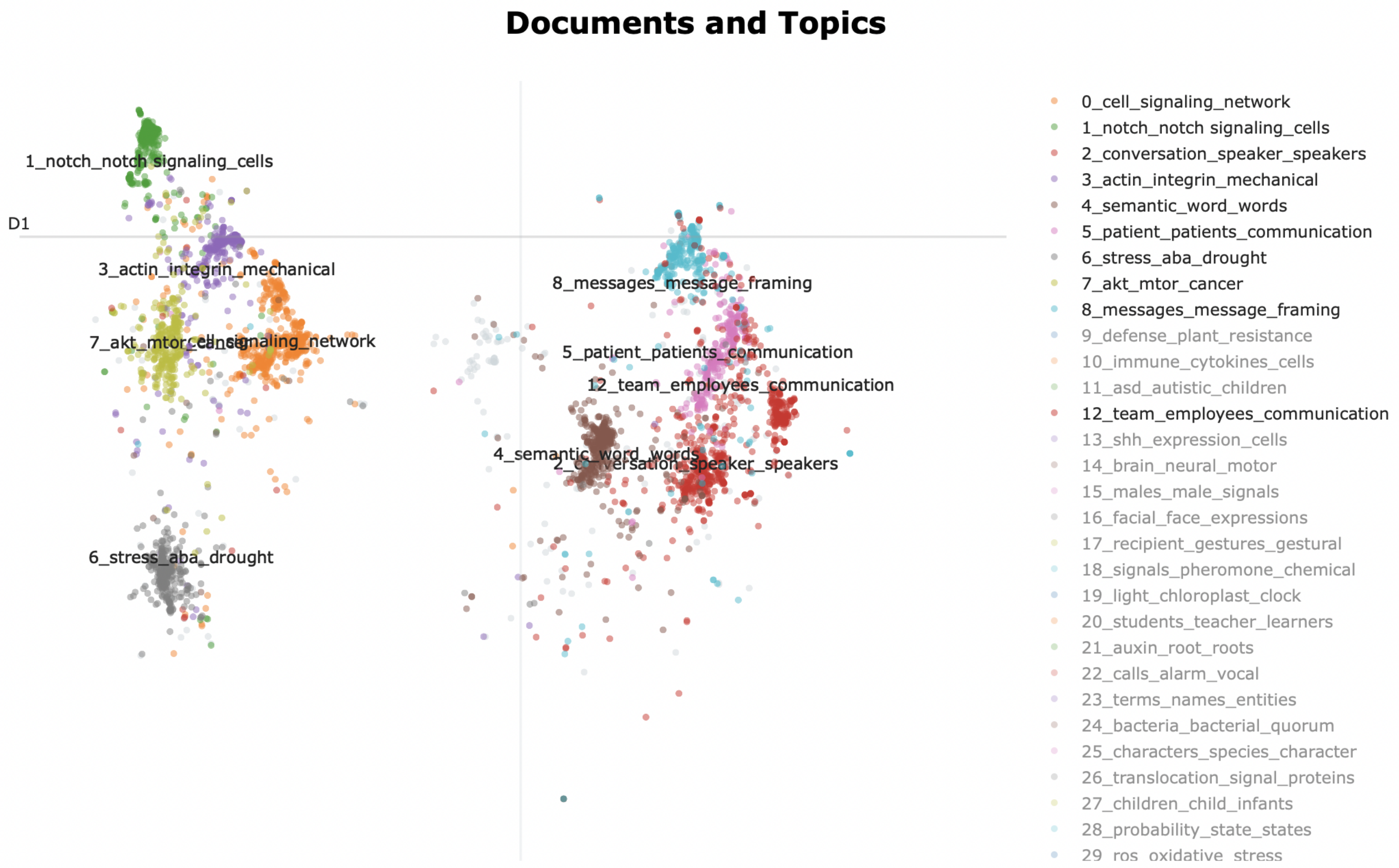


Figure 5.1: UMAP visualization of the paragraph model results (107 topics in total), with a subset of 10 most popular topics visible for clarity—selected topics are representative of the two "outer" clusters—distinguishing between cellular signalling and social communication. Each point represents a 2D-reduced vector embedding of a document (paragraph). Colors correspond to the topics. See text for description.

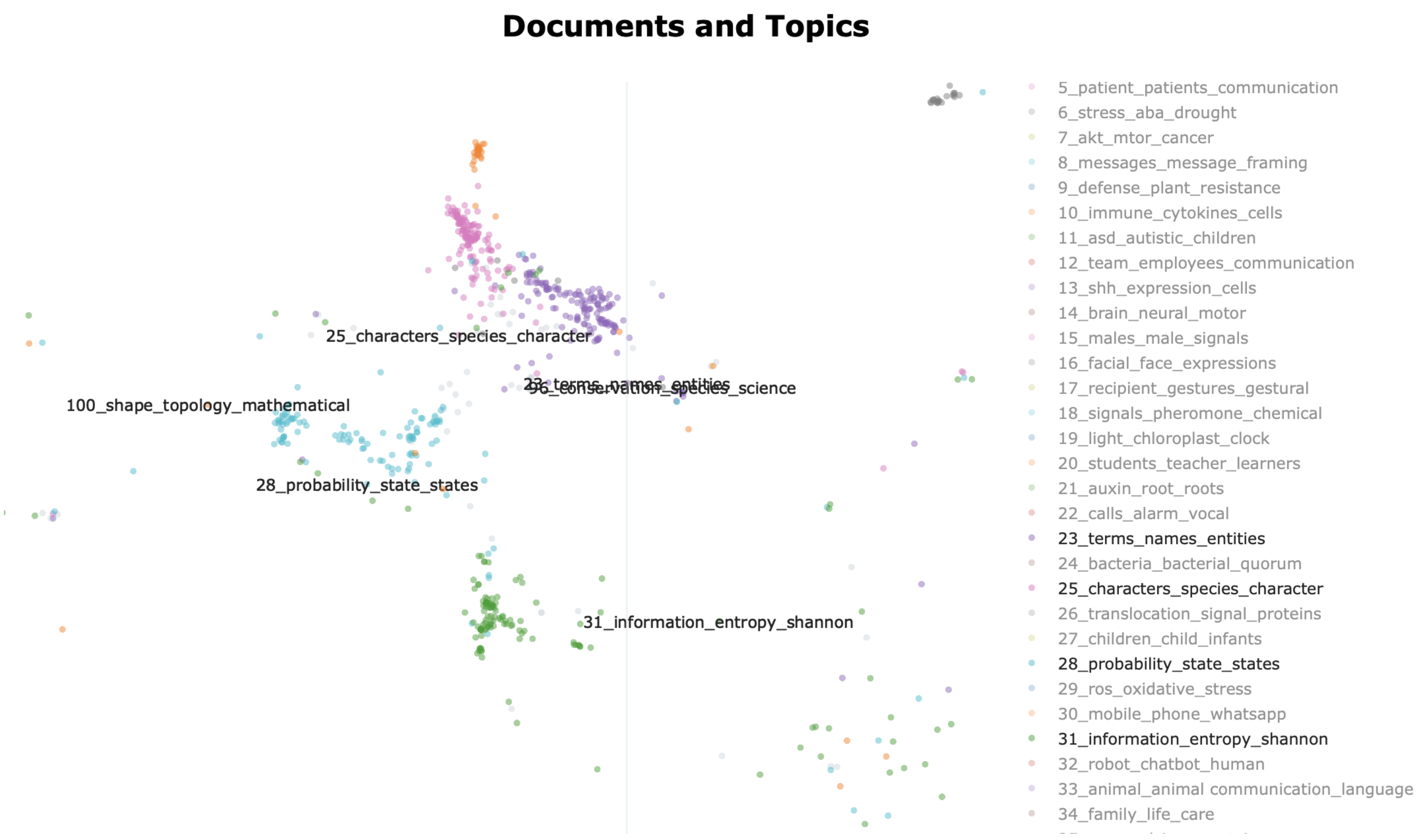


Figure 5.2: UMAP visualization of the paragraph model results (107 topics in total), with a subset of topics visible for clarity—selected only topics that form the "inner" cluster, covering "meta" references to meaning. Each point represents a 2D-reduced vector embedding of a document (paragraph). Colors correspond to the topics. See text for description. Note that the scale is different than in Figure 5.1.

e.g. “communication at the end of life potentially has the greatest and longest-lasting impact on family members” (1840600, 77).

A few topics refer to “meaning” directly (i.e., “meaning” is one of the keywords for the topic; these are mostly the “higher-level”—social—topics):

- topic 2 is concerned with interactive character of meaning,
- topic 4 with the processing of word meaning in the brain,
- topic 27 with language acquisition,
- topic 38 with how humans express semantics outside of language—in art, especially music and visual arts,
- topic 45 explores the different mappings between signs and their reference in sign language,
- topic 58 focuses on how meaning is modelled in sender/receiver models of communication,
- topic 69 models molecular processes as encoding/decoding,
- while topic 88 concerns the relation between syntax and meaning.

Some topics refer to meaning indirectly, for example suggesting what the communication might be about: for example, key terms of topic 6 include “stress” and “drought”. However, the ascription of semantics is not done directly: “Other plant phytohormones [...] act as secondary players and modulators of main signaling pathways. For instance, ABA [abscisic acid] has a primary role in the regulation of plant defenses against abiotic stressors” (15627565, 173). The example indicates that ABA signalling is initiated in response to abiotic stress—suggesting that those signals refer to or are about those stressors, such as drought, and is a part of paragraph that was picked out by semantic search through the sentence “Interaction between individual components of a highly sophisticated network of signal transduction pathways can be additive, antagonistic, or synergistic.” which is averagely related to several queries, most closely to Shea’s varitel definition—providing some confidence into this semantics-oriented interpretation.

Results model

To pursue this further, I’ve trained a topic model directly on individual sentences produced by semantic search (I’ll refer to it as the “results” or “sentence-level” topic model, to distinguish it from the “paragraph-level” model just discussed)[8]. The results model (with the default minimum cluster size of 50) picked out 25 topics with an unequal distribution of individual documents, split into two larger clusters—as illustrated in Figure 5.3. One of the clusters is dominated by the topic 0, COMMUNICATION INFORMATION, to which over half of the documents

---

[8] I have also trained another sentence-level model on the whole paragraphs that were produced by semantic search (sent-par model). However, this model produced topics with very low coherence scores even at high granularities and as such was not reliably informative and will not be analyzed here. The model and visualizations produced are included in online supplementary materials, available at: https://doi.org/10.17605/osf.io/ax4jm.

were ascribed. Interestingly, this topic is significantly skewed towards the results most similar to the queries used. The average highest similarity (among queries) is approx. 0.43, significantly above other topics, which are in range of $0.35 - 0.37$ (the second topic by this criterion is topic 22, GOSSIP, with the average highest similarity of approx. 0.39, which belongs to the same cluster). Upon a closer inspection, topic 0 is a bit of an amalgamate: it includes references to human linguistic communication, to other systems of animal signalling, a variety of uses of information theory—either as a theoretical frame, or as a formal, mathematical model, and some mentions of neuronal processes, e.g.:

> "We found that the representational pattern of a PE reliably predicts the identity of the agent to whom the signal is attributed, consistent with a neural self-other distinction implemented via agent-specific learning signals." (13714782, 6)

> "Input specificity refers to the observation that when a tetanus (or another induction protocol) is applied to one of two (or more) inputs converging on the same post-synaptic target then the potentiation is only observed at the tetanised input." (16990635, 29)

What seems to be dominant across those different contexts, is the reliance on the notion of information—not only in the specific sense motivated by information theory (as discussed in chapter 4), but more broadly, e.g.:

> "These signals deliver information about kinship, immunogenetic mating characteristics, the endocrine status, and emotional states reflecting potential transient danger." (18782890, 102)

Information recurs as a way to conceptualize what happens in communication across the different contexts captured by this topic. This topic is accompanied only by the relatively small topic 22, GOSSIP, which clusters together precisely that—mentions of gossip as the topic of psychological research (semantic properties are highlighted by the consideration of honesty and the inclusion of value judgments).

The second cluster includes the remaining 23 topics. It is dominated by topic 1, CELL SIGNALLING, which includes a variety of cellular signalling processes, their molecular characteristics, and (computational) models. All other topics in this cluster are also placed at the molecular scale and significantly less frequent. These are focused primarily on specific signalling structures, e.g., the Notch pathway (topic 2), calcium channels (topic 3), ABA (topic 5) or cytokines (topic 6).

This distinction is further exemplified by the topic similarities (Figure 5.4): while there are very few strong similarities between topics (despite a relatively high baseline similarity, which is expected, considering the narrow focus of the queries), they go along the division into two clusters, with topic 0 being more similar only to topic 22, and topic 1 to almost all others.

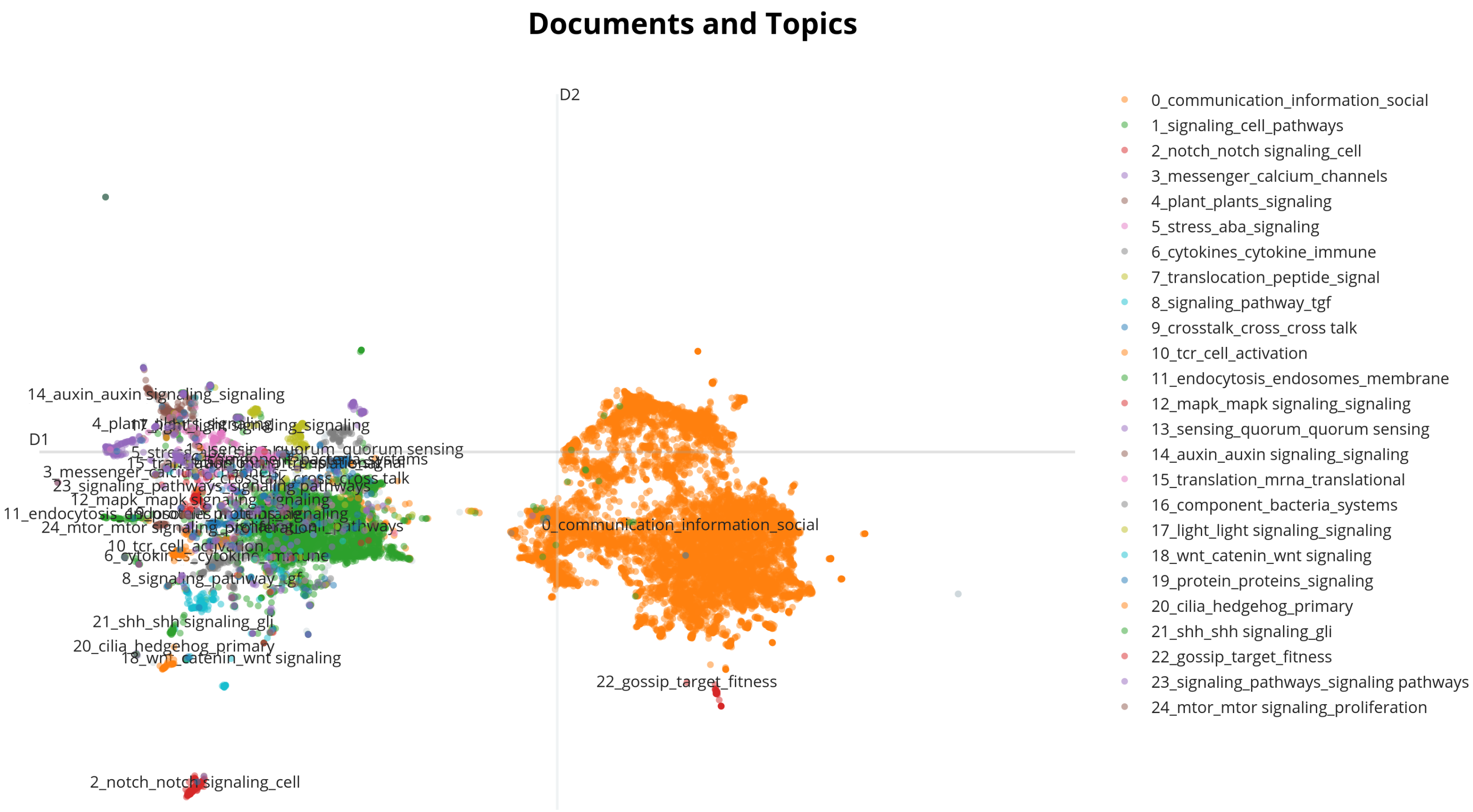


Figure 5.3: Two-dimensional representation of document vector embeddings, calculated by BERTopic using UMAP algorithm. The documents are clustered by topics based on the results topic model. Each point represents a 2D-reduced vector embedding of a document (paragraph). Colors correspond to the topics. The full, interactive version is available in the online supplement, available at: https://doi.org/10.17605/osf.io/ax4jm.

Figure 5.4: Topic similarities heatmap for the results topic model (25 topics). Topic similarities are calculated by the BERTopic algorithm as the cosine similarity between vector embeddings of topics.

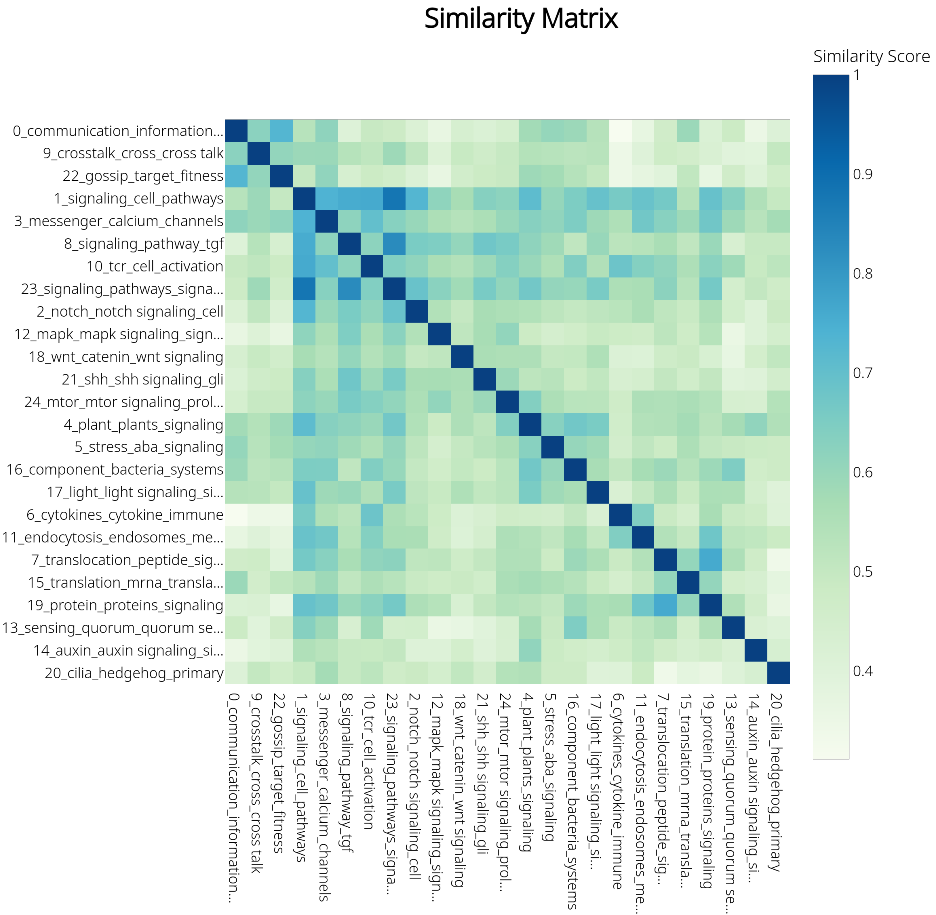


Table 5.2: 10 most frequent topics with document counts for the results model, excluding the “outliers” topic which was ranked 3rd overall (577 documents). I have chosen the topic names through the qualitative review of the different representations of the topics produced by BERTopic and a close reading of the representative documents selected by the algorithm (see chapter 2). The table includes all representations used, limited to top 5 key terms for each. Base representation is the default algorithm, the c-TF-IDF (continuous term frequency—inverse document frequency, see subsection 2.2.3). “Nouns and modifiers” and “Verbs and modifiers” use Spacy part-of-speech tagger with the `en_core_web_sm` model to extract terms from the c-TF-IDF which match specified syntactic patterns—in this case, nouns or verbs and their modifiers (if any). “KeyBERT + MMR” calculates the vector distance (semantic similarity) between candidate keywords and topic embeddings, selects best candidates (least distance), and filters them with the Maximal Marginal Relevance algorithm, to minimize redundancy between keywords. Full list of topics is available in the online supplement at: https://doi.org/10.17605/osf.io/ax4jm.

Topic	Count	Name	Base representation	Nouns and modifiers	KeyBERT + MMR	Verbs and modifiers
0	9187	COMMUNICATION INFORMATION	communication, information, social, language, conversation	communication, information, social, language, conversation	communicating, conversations, communicative, conversational, verbal	used, use, communicate, based, using
1	3385	CELL SIGNALLING	signaling, cell, pathways, cellular, signal	signaling, cell, pathways, cellular, signal	signaling pathways, cell signaling, signalling pathways, signaling pathway, cellular signaling	signaling, signalling, signals, function, involved

Continued on next page

Table 5.3: (continued)

Topic	Count	Name	Base representation	Nouns and modifiers	KeyBERT + MMR	Verbs and modifiers
3	313	CALCIUM SIGNALLING	calcium, messenger, signaling, second messenger, channels	calcium, messenger, signaling, second messenger, channels	calcium signaling, signaling pathways, cellular signaling, signalling, downstream signaling	signaling, signals, involved, mediated, sensed
2	301	NOTCH SIGNALLING	notch, cell, signaling, pathway, cells	notch, cell, signaling, pathway, cells	notch signaling, signaling notch, pathways notch, notch, signaling pathway	signaling, conserved, signalling, involved, plays
5	250	STRESS	stress, stomatal, responses, abiotic, pathways	stress, stomatal, responses, abiotic, pathways	aba signaling, stress signaling, signaling pathway, abiotic stress, cellular stress	signaling, signal, involved, signalling, mediated
4	232	PLANT SIGNALLING	plant, plants, signaling, development, rice	plant, plants, signaling, development, rice	plant signaling, signaling plants, plant hormones, plant hormone, signaling pathways	signaling, involved, molecules, mediated, play

Continued on next page

Table 5.3: (continued)

Topic	Count	Name	Base representation	Nouns and modifiers	KeyBERT + MMR	Verbs and modifiers
23	206	SIGNALLING PATHWAYS (GENES)	signaling, pathways, pathway, complex, accumulation	signaling, pathways, pathway, complex, accumulation	signaling pathway, components signaling, transcription factor, scf coi1, genes	signaling, mediated, interact, play, involved
9	196	CROSSTALK	crosstalk, cross, talk, pathways, signaling	crosstalk, cross, talk, pathways, signaling	pathways crosstalk, signaling crosstalk, signaling cross, talk signaling, signalling pathways	crosstalk, talk, signaling, referred, occur
6	192	CYTOKINES	cytokines, cytokine, immune, cytokinin, cells	cytokines, cytokine, immune, cytokinin, cells	cytokine signaling, cytokine receptor, cytokine receptors, cytokines, cytokine cytokine	cytokine, signaling, receptor, activated, bound
7	186	SIGNAL PEPTIDES	peptide, translocation, signal, peptides, protein	peptide, translocation, signal, peptides, protein	signal peptide, signaling peptides, proteins signal, secreted proteins, secretory pathway	signal, membrane, targeting, secreted, predict

The emerging picture suggests that there is a discontinuity in how the contents of the signals are individuated in different areas. While information emerges as dominant in the context of organismic (or social) interactions—both linguistic and non-linguistic—the smaller spatiotemporal scales of biological organization seem to be focused more on the materiality of communicative processes. Even when the passages discuss specific computational models, as in the following example: “A feedback signal from the Q-value unit selected for action creates synaptic tags on synapses that carry a synaptic trace (step iv).” (16554868, 58), the focus is on elements such as “synaptic tag” or “synaptic trace” which are simulations of how neurotransmitters can act on synaptic connections to strengthen them. Considering these sentences were picked out by querying for ascriptions of meaning to signals, we may take that to mean that in the cellular context, the meaning is specified by the particular material makeup of the communicative process, while on the larger spatiotemporal scale, the meaning emerges from an informational organization of the process.

5.3.2 *Word Sketch*

The exploration of Word Sketches also provides somewhat limited insights in this context: as visible from the examples discussed above, there are no clear and consistent textual markers of ascriptions of meaning identified through semantic search. Verbs “mean” and “refer”, “meaning” as a noun, and “referential” as an adjective appear relatively frequently. Interestingly, they seem to be more closely connected to the cell-level cluster of topics, despite more frequently appearing in the organism-level one. Comparing the cosine similarity of these words to topic embeddings (as previously discussed in section 3.4—since BERTopic produces topic embeddings, embeddings of individual terms can be directly compared; this is noted as S_c) in the results topic model, it emerges that “mean” and “meaning” are most similar to topic 1, CELL SIGNALLING (“mean”: $S_c \approx 0.48$, “meaning”: $S_c \approx 0.49$), followed by topics 6, CYTOKINES (“mean”: $S_c \approx 0.43$, “meaning”: $S_c \approx 0.36$), and 16, TWO-COMPONENT SYSTEMS (“mean”: $S_c \approx 0.43$, “meaning”: $S_c \approx 0.41$). In the case of “refer”, this is inverted, with topic 16 being the most similar ($S_c \approx 0.44$), followed by 23, SIGNALLING PATHWAYS (GENES) ($S_c \approx 0.43$), and then topic 1 ($S_c \approx 0.42$), and for “referential” the similarities to 1 ($S_c \approx 0.43$) and 16 ($S_c \approx 0.41$) dominate. Topic 0, COMMUNICATION INFORMATION, is significantly less similar, appearing as 5th most similar for “referential” and only as 10th and 11th for “mean” and “meaning”, respectively (for “refer” the similarity is lower than the algorithmic threshold for the relevant BERTopic function and the exact value was not calculated).

Importantly, in case of “mean” this is largely driven by the sense of “average” or by the sense of “through” (i.e., “by means of”), and is entirely irrelevant to the question of semantics, e.g.:

> “The signaling vector was defined as the vector sum of the vectors pointing from the cell centroid to the center of each included window, weighted by the FRET ratio for that window minus the mean FRET ratio over all windows.” (2040833, 201)

“Refer”, on the other hand, appears primarily in a “meta” sense when terms used are being defined or operationalized (e.g., “From here onwards, we refer to this SuH dependent N activity that promotes expression of E(spl)C RNA as SuH/N intra signaling.” (9391147, 21)). Finally, “referential” has only a single (somewhat underspecified) use not in the context of topic 0: “Our findings are also relevant to our knowledge of functionally referential signalling.” (topic 1, CELL SIGNALLING; 17383402, 159). “Meaning”, on the other hand, while predominantly used in the context of topic 0 (despite low cosine similarity, 532 out of 536 uses of the string “meaning” in the semantic search results were classified as belonging to topic 0), has some important contexts in other topics that have semantic relevance (see box 5.3).

With these limitations in mind, Word Sketch results still provide some important insights. For “meaning” either as a verb form or a noun, modifiers indicate what is the vehicle or source of semantics: “word” or “utterance” meanings are accompanied by “actor” or “speaker” meanings. These can be “intended”, “conveyed”, “encoded”, “inferred”, “implied”, “expressed” and “accessed” (these are all verbs that use “meaning” as an object). Further, they indicate how meanings are established, as they can be “conventional”, “dictionary” or “archetypal”, as well as “literal” and “metaphoric”, “implicit” or “dominant”, “subordinate”, “structured” or simply—“wrong”. “Meanings” have a purpose: they “bestow”, “trigger” or “transform”, and are a dynamic process, as they can “emerge”, “become”, “arise” (all verbs with “meaning” as a subject). At the same time they are inherently structured: they “correspond”, and are listed in the same breath as “form”, “association”, “structure”, “order”, or a “mapping”. Note however, that this is primarily driven by the organism-level[9] context and included in topic 0 (see box 5.4 for examples which especially highlight the structural properties of meaning).

---

[9] For the sake of stylistic clarity, I use the notion of “organism-level” to refer to forms of communication exhibited by multicellular organisms, including humans, non-human animals and plants, including their social interactions, and the notion of “cell-level” to refer to forms of communication exhibited by individual cells (whether they form part of colonies, multicellular bodies, or are free-living unicellular organisms) and tissues (groups of cells embedded in a larger organism). While this may not be the most precise distinction, I hope the intuition behind it, motivated by empirical data—the clustering in BERTopic results, is clear enough.

Box 5.3: All uses of the word “meaning” in semantic search results which were not classified as topic COMMUNICATION INFORMATION (results model).

Most of these sentences were produced by several queries, here they are classified with the best query (the full list with scores can be found in the online supplement, available at: https://doi.org/10.17605/osf.io/ax4jm). The similarity scores are rounded up to three decimal places. The reference includes the CorpusID from S2ORC and the sentence number.

1. SENTENCE: Specifically, we discuss their meaning, the consistency among different network inference methods, ensemble methods, the assessment of GRNs, the estimated number of existing GRNs and their usage in different application domains.
   ID: 693936, 3
   AREAS: developmental biology
   TOPIC: 1 (CELL SIGNALLING)
   BEST QUERY: 6
   SIMILARITY TO BEST QUERY: ≈ 0.353

2. SENTENCE: Both systems have been extensively used for Ca 2+ -signaling research, but the meaning of their Ca 2+ signatures remains elusive.
   ID: 51934500, 51
   AREAS: plant science
   TOPIC: 3 (CALCIUM SIGNALLING)
   BEST QUERY: 2
   SIMILARITY TO BEST QUERY: ≈ 0.359

3. SENTENCE: Numerous distinct signal transduction systems have evolved to allow cells to receive these inputs, to translate their codes and, subsequently, to expand and integrate their meanings.
   ID: 52895117, 1
   AREAS: plant science
   TOPIC: 1 (CELL SIGNALLING)
   BEST QUERY: 2
   SIMILARITY TO BEST QUERY: ≈ 0.370

4. SENTENCE: Many forms of CWD, both abiotic and biotic, activate common signaling processes involving hormone crosstalk, ROS and calcium signaling, meaning that the transcriptional machinery, controlled by these signaling processes, will be activated in response to CWD as well (Le Gall et al. 2015;Li et al. 2016a, b, c;Liao et al. 2017;Nováković et al. 2018).
   ID: 249464716, 181
   AREAS: molecular biology, plant science
   TOPIC: 9 (PATHWAYS CROSSTALK)
   BEST QUERY: 2
   SIMILARITY TO BEST QUERY: ≈ 0.344

Box 5.4: Manually selected examples of the use of the word "meaning" in topic COMMUNICATION INFORMATION (results model).

Most of these sentences were produced by several queries, here they are classified with the best query (the full list with scores can be found in the online supplement, available at: https://doi.org/10.17605/osf.io/ax4jm). The similarity scores are rounded up to three decimal places. The reference includes the CorpusID from S2ORC and the sentence number.

1. SENTENCE: Some signals, for example, can be honest intrinsic correlates of physical and genetic attributes of senders (so called indices) in which a signal's meaning is associated with its method of production.
   ID: 1164365, 35
   AREAS: Animal behaviour, Ecology and evolution
   TOPIC: 0 (COMMUNICATION INFORMATION)
   BEST QUERY: 2
   SIMILARITY TO BEST QUERY: ≈ 0.328

2. SENTENCE: The evolution of this type of signaling system has been modeled by Ay et al. (2007) in the context of robustness of signal design: a signal is robust if it can withstand partial occlusion, which it can if it contains separate modular clusters that are at least to some extent correlated in meaning.
   ID: 1855883, 34
   AREAS: Animal behaviour, Ecology and evolution
   TOPIC: 0
   BEST QUERY: 2
   SIMILARITY TO BEST QUERY: ≈ 0.347

3. SENTENCE: To date, the study of visual, non-verbal social interaction has generally focused on investigating the meaning of the discrete behavioral actions and gestures as sources of communication (e.g., Wosegien and Lamprecht 1989).
   ID: 3569830, 15
   AREAS: Animal behaviour
   TOPIC: 0
   BEST QUERY: 3
   SIMILARITY TO BEST QUERY: ≈ 0.360

4. SENTENCE: An inter-signal interaction hypothesis can thus either be efficacy-based (one signal enhances the detection or processing of another) or content-based (the meaning of one signal depends upon the presence of a second signal).Since every signal is under selection for both efficacy and content, explanations from each of these perspectives are not mutually exclusive.
   ID: 5038555, 66
   AREAS: Animal behaviour
   TOPIC: 0
   BEST QUERY: 1
   SIMILARITY TO BEST QUERY: ≈ 0.372

5. SENTENCE: Assessment of the words positively or negatively associated with each dimension (Table 2) demonstrates the semantic coherence generated across observers (i.e.

that the different individual observers use terms with similar meanings when viewing particular behavioural expressions).
ID: 7994118, 119
Areas: Animal behaviour
Topic: 0
Best query: 6
Similarity to best query: ≈ 0.375

5.4 CONSIDERING ORGANISM-LEVEL MEANING

The appearance of “mapping” in Word Sketch results is quite telling. In philosophy of mind, some theoreticians distinguish a tradition within the representationalist approach to cognition that is often termed “encodingism” (this view is classically ascribed to approaches such as the language of thought hypothesis, Fodor 1975; Fodor 1992; Fodor 1998; the term has been coined by Bickhard and Terveen 1996). The “encodingist” paradigm asssumes that there are symbolic mental representations which correspond to (encode)[10] world elements. This view is quite widespread, with many applications in neuroscience (e.g., A. K. Robinson, Quek, and Carlson (2023); see the overview and criticism in Brette (2019); and commentaries to that article). From the perspective of naturalizing semantics, the encodingist approach usually does not specify a particular source of the correspondence. However, it significantly constrains the structure of an explanation for that correspondence. In many forms of this theoretical perspective, there are additional restrictions on that correspondence relation: e.g., it is presumed to be an isomorphism.

This approach is heavily criticized especially by the proponents of various flavours of antirepresentationalism. Conceptually, antirepresentationalists argue against the existence of representational vehicles (e.g., Chemero 2000, 2009), the possibility of clearly and objectively individuating referents in the world (e.g., Gibson 1960), or by highlighting the “homunculus” inherent in this view (Bickhard and Terveen 1996). On factual grounds, they point out, among others, the static character of encodingist models (their lack of temporal dimension) and resulting limited explanatory power (see Bickhard and Terveen 1996).

Based on those criticisms, ecological psychologists try to develop alternatives for the encodingist approach, such as Mark Bickhard’s interactivist model (Bickhard 2007, 2009, 2020), Terrence Deacon’s “Symbolic Species” approach (Deacon 1998, 2011, 2021), Howard H. Pattee’s work (Pattee 1997; Pattee [1969] 2012; Pattee [1982] 2012; elaborated by Rączaszek-Leonardi 2012; Rączaszek-Leonardi 2016), or the emerging resonance models of the central nervous system (e.g., Raja 2018; Raja 2021; Raja 2024; see also Falandays, Yoshimi, et al. 2023). These alternatives will be important further down the line and I will discuss them in greater detail in the next section. However, it needs to be noted already at this point that these organism-level considerations have a complex relation to the debates of naturalization of semantics discussed in the beginning of the chapter. With that in mind, I bracket this topic for now and I will get back to it towards the end of this chapter.

---

[10] Note that this is a significantly different notion of correspondence compared to the one that Miłkowski’s correspondence theory invokes: standard encodingist views, that Bickhard argues against, consider correspondence as a formal mapping of the states of representational system to the (formal categorization of) states of the world, not in terms of satisfaction conditions. In fact, Miłkowski (2025) considers the relation of the correspondenceview of semantic information to Bickhard’s interactivism in detail.

Despite these criticisms, the broadly encodingist approach remains dominant in cognitive science. The relevance of notions such as "mapping", "correspondence", or "association" in the empirical results discussed just above highlights this fact. Consider examples from box 5.4. Example 1, in stating that "signals […] can be […] correlates of physical and genetic attributes of senders" and that "signal's meaning is associated with its method of production", highlights how signals can be individualized. Example 3 takes "discrete behavioral actions" to be meaningful. Example 4 in postulating that signals are "under selection for […] content" assumes that the content is a stable feature. Finally, example 5 postulates a correspondence, and a socially shared one at that: "the different individual observers use terms with similar meanings when viewing particular behavioural expressions", which presumes a mapping between terms and viewed expressions. While neither of these examples is overly restrictive in the kinds of relationships it admits ("correlated" and "associated" suggests a relatively weak relationship; "similar meanings" precludes a strict one-to-one mapping), they are all clearly embedded in the encodingist perspective.

However, this correspondence is not the primary factor which comes out of the examples. Consider the sentence:

> "In other words, when communicating, the words have the primary role of expressing, on one side, a set of experiences that the utterer wants to focus on or convey, and of raising, on the other (or on the reader's side), an analogous set of personal experiences." (342700, 113)

While an encodingist approach is clearly central to understanding how experiences are conveyed by the words, a *pragmatic* function of the exchange comes to the fore. Likely because of the focus on communication, with its inherent distinction between a sender and a receiver (see chapter 3), meanings are not treated as something that just "is". The utterer's intentions and reader's interpretation is crucial. Meaning becomes an act. This is clearly highlighted in the Word Sketch results indicating the dynamic character of meaning: how it must emerge, arise, become or how it must be intended, implied, expressed, as well as inferred and accessed. This highlights that communication is primarily a pragmatic circumstance. Since semantics and pragmatics are likely distinct, yet deeply intertwined (e.g., McNally 2013), this inherent dynamics of use introduce important constraints on the conceptualization of meaning. We can explore them in two philosophical directions: through relevance theory and the understanding-first approach.

5.4.1 *Intention and relevance*

Paul Grice's influential work (Grice 1957, 1969, 1995) and the framework of relevance theory, expanding on the Gricean approach (Sperber and D. Wilson

1995; D. Wilson and Sperber 2010) aims to capture precisely this dynamic between semantics and pragmatics of communicative processes. In this context, the approach developed by Thom Scott-Phillips with Christophe Heintz is quite significant, as it focuses explicitly on applying the relevance theory framework to animal signalling and the explanation of the evolution of human language (Scott-Phillips 2008; Scott-Phillips, Kirby, and G. R. Ritchie 2009; Scott-Phillips 2025; Heintz and Scott-Phillips 2023; Scott-Phillips and Heintz 2023).

Scott-Phillips and Heintz aim to underscore the role of pragmatics and their cognitive underpinnings in this transition. They do so by centering on the concept of "intention" (also visible in the empirical sample), crucial to relevance theory. "Intention" refers to a mental state motivating behaviour. In the context of communication, the framework distinguishes "informative and communicative intentions, which are proximate, cognitive processes for the functional tasks of expressing and communicating, respectively" (Heintz and Scott-Phillips 2023, p. 4). Informative intentions underlie actions "that intentionally change mental states, and which can do so without overtly bringing attention to the intention itself" (Heintz and Scott-Phillips 2023, p. 4). An example of such an action Heintz and Scott-Phillips provide is dressing in a particular way for a job interview to signal our competence.

Communicative intentions are a subset of informative intentions, as they motivate behaviours aiming to inform an audience but also "to make the actor's informative intention mutually known" (Heintz and Scott-Phillips 2023, p. 5). Here, an example they provide is eating berries next to somebody to signal their edibility (which is an action driven by informative intention), compared with a gestural representation, miming of eating the same berries to signal their edibility—which draws attention to the signal itself, and hence is an action motivated by communicative intention. Importantly, for Heintz and Scott-Phillips, sender's intentions are not enough: the individual must be also embedded in a particular social structure which enables the audiences to interpret those intentions—whether overt or not—from the actions.

In some cases—where a particular social ecology with a selective pressure for cooperation persists on an evolutionary timescale—the cognitive capacities for producing and interpreting informative intention can co-evolve, eventually leading to the emergence of ostensive[11] communication and the generative "unleashing" of expression that in their account sits at the core of linguistic productivity. While they argue that ultimately in the course of evolutionary history on Earth these circumstances have aligned only for humans, some animals might be capable of "proto-ostensivity" in their view, whereby through "non-standard life histories", in particular human impact on rearing conditions (see Call 2011),

---

[11] While the term "ostension" functions in the literature in several meanings, Heintz and Scott-Phillips use it to refer to cases of communication where the communicator "provides evidence for the informative intention itself", making the exchange "overtly intentional" (Heintz and Scott-Phillips 2023, p. 5).

a " 'proto' presumption of relevance could result [...] in non-human apes" (Heintz and Scott-Phillips 2023, p. 10).

This view accomplishes two things: it highlights the continuity of the evolutionary trajectory of human language, at the same time clearly identifying the point of transition. While there are some important limitations to Heintz and Scott-Phillips' view, in particular with their treatment of development and the adaptationist focus of the proposed account (see Rorot et al. 2022), what is particularly relevant in the context of empirical results discussed in the previous section is that even before the transition to "unleashed" expression they closely tie the meaning of communicative actions to communicator's intentions. This, more broadly, is a view that several proponents of relevance theory have explored as an account of semantics (rather than pragmatics, as has become standard to interpret Gricean concern with speaker's meaning in other lines of inquiry in philosophy of language; note that relevance theory on its own focuses on linguistic, most often verbal, communication).

While Heintz and Scott-Phillips do not explicitly consider how meanings of signals are established, other philosophers have advanced possible explanations within the relevance theory framework. Emma Borg (Borg 2009) distinguishes two approaches: one more directly inspired by Grice (1995), what she terms A-style intention-based semantics, and one emerging from Sperber and Wilson's work (1995), termed B-style intention-based semantics. The difference between the two views is that in the A-style version, speaker's intention plays a more minimal, preconditional role, couched in "a convention among a community of speakers to use an expression of type x in the way specified by the given instance of [utterer's meaning]" (Borg 2009, section 12.1), and the hearers need not have access to the mental states of the actual, current speaker.

Meanwhile, in the B-style intention-based semantics a speaker must engage in inferential reasoning taking the utterance heard as a piece of evidence with regard to the "machinations of relevance, which in turn serve to make speaker intentions evident" (Borg 2009, section 12.2). Note that this distinction at least superficially tracks the difference between "informative" and "communicative" intentions outlined above. Borg argues for A-style intention-based semantics as more accurate for theories of linguistic meaning, with the corollary argument that linguistic meaning bears at most a family resemblance to general ostensive behaviour and should be analyzed as involving importantly different comprehension processes. She claims that "[t]he meaning of a non-linguistic ostensive act seems ineliminably tied to its context in a way that the meaning of a linguistic act is not" (Borg 2009, section 12.4): the inclusion of linguistic convention allows the meaning to detach itself from immediate context, including current speaker's intentions (this approach strongly echoes Deacon's "Symbolic Species" hypothesis and his account of Peircean semiotics that I'll discuss in greater detail below, in subsection 5.5.1).

This is in clear tension with Heintz & Scott-Phillips' account, as they posit language as a species of ostensive communication more broadly, exploiting the cognitive capacities required for unleashed expression (Heintz and Scott-Phillips 2023, p. 13-14). At the same time, they do point out conventions as a distinguishing feature, identifying their source in the replicability of behaviours (note that this relates closely to Pattee's views, which will be discussed in greater detail below in subsection 5.5.1, see also Rączaszek-Leonardi 2012).[12]

Hence, from the perspective of relevance theory, or at least the version Heintz and Scott-Phillips advocate for, we can view organism-level meanings, whether they concern linguistic or non-linguistic forms of communication, as directly tied to speakers' and hearers' activities and their mutual entwinement. At the same time they remain heavily structured, and while the particular mapping of signals to the world is actively constructed (on the basis of extant systems), the view advanced by relevance theory retains some "encodingist" elements (in particular as, at least according to Ruth Kempson, it remains embedded in a Fodorian version of the representational theory of mind, see Kempson 1990, p. 16). But it complicates the encodingist picture significantly, especially through the introduction of the dynamic, contextual elements and the Gricean treatment of communication pragmatics.

5.4.2 *Understanding and meaninglessness*

On the other hand, we can consider the dynamics of use visible in empirical data in the context of Niklas Dahl's "understanding-first" approach to communication (Dahl, forthcoming). Dahl highlights that discussions of semantics often simply gloss over the concept of UNDERSTANDING taking it to be a "right relationship to meaning" (Dahl, forthcoming, p. 1), despite its role as a criterion of communicative success. He draws on the epistemological literature on understanding to develop a more nuanced notion of linguistic understanding, boiling down to two components: "(i) Knowing how to use an expression correctly. (ii) Having cognitive recognition of that knowledge how." (Dahl, forthcoming, p. 11). While his view is positioned squarely within contemporary debates in philosophy of language, here I want to explore it slightly beyond its intended scope, as the idea of "rejecting the picture of communication as the transmission of meanings or references" (Dahl, forthcoming, p. 29), while somewhat at odds with my own understanding of communication developed in chapter 3 (though perhaps superficially), provides an enlightening perspective for the current considerations.

---

[12] In fact, Scott-Phillips and Heintz (2023, p. 104-106) develop a more complex taxonomy of behaviours intended to manipulate others' attention. In that account they ascribe to non-human great apes what they term "Ladyginian communication": intentional manipulation of attention towards one's own intentions. They contrast it with "Gricean communication", which is aimed specifically towards one's own *informative* intentions. But Gricean communication is not the end of story: it has a further subset of "Lewisian communication", which introduces convention, and if those conventions form a structured network, it becomes "Saussurian communication". While this model is not central to the current analysis, it is important to note that it doubles down on this "layered" view of ostensive communication, with meaning appearing relatively early on.

In Dahl's view, "meaning-talk" emerges from conversational repair (a point he draws from Gauker 2015; this is in fact topic widely explored in cognitive science, e.g., Dingemanse et al. 2015, though neither Dahl nor Gauker make an explicit connection) and is—or at least initially was—limited to relating the expression in question to other expressions that we are already familiar with. While this deflationary notion of "meaning" has some philosophical limitations (Dahl, forthcoming, p. 27), it helps us make sense of several uses of the notion in the corpus. The sentence "During communication, we observe and evaluate the recipient's feedback to adjust the content of the output message." (257137275, 131; area: psychology) confirms Dahl's recognition of understanding as crucial for the success of communicative exchanges. Meanwhile, the sentence "In humans and honey bees, the abilities of the listener directly constrain the meaning of information exchanged during an interaction." (227241123, 108; areas: psychology, ecology and evolution): while we can look at it through the relevance-theoretic lens of what informative intentions the listener (in the bees' case more precisely "the toucher") is able to infer, the understanding-first approach is even more illuminating, as it highlights the limits of recipients' understanding resulting from the exchange. Similarly, in the sentence:

> "More specifically, it is the outcome of communication operations which, through redundancy and repetition, perform a continuous and selective re-imprinting of meaning that can be used for communication." (264054924, 4; Area: psychology),

while we can consider it as a reference to the emergence of convention, highlighted by relevance theorists, Dahl's approach (embedded in the tradition of Wittgenstein 1968 and Michael Dummett 1995, see Dahl, forthcoming, pp. 5-6) of relating understanding to the knowledge how to use an expression, helps view this as a description of how the understanding of "communication operations" is acquired and established at the same time, accounting for the dynamics of semantics coming out of empirical results. Finally, Dahl's model helps acccount for the prevalence of "understanding" in semantic search results, as it appears quite frequently, with 446 direct mentions of "understanding", out of which 296 appear in the context of topic 0, COMMUNICATION INFORMATION (with the remaining mentions assigned to other topics being almost exclusively in reference to the scientific understanding of cell-level communicative processes).

Importantly—as we saw already in the discussions of the role of communication in chapter 3—there's an aspect of goal-directedness which the understanding-first approach captures through the notion of linguistic function which originated in classical structural linguistics (Bühler [1934] 2011; Jakobson [1960] 1981; although Dahl, forthcoming, does not cite directly the originators of this approach). The goal-directedness underscores normativity of meaning which depends on the *goal* of communication. The normativity, in turn, highlight that we should look for the purpose of the communicative instances when analyzing the data, and the standard catalogues of linguistic functions help guide our attention (e.g., Jakobson ([1960] 1981); see also the discussion in the chapter 1 which reviewed Jakobson views on relations between linguistics and biology).

Dahl's view is clearly not an appropriate description of how "meaning" is used by scientists (and is not intended as such). There are many mentions of communication as the "conveying of meanings", and a relevance theory seems to be well-established in this context, e.g.:

> "Most human communication is a compound of coded conventional symbolic meaning (as in language), and inferences about communicator intent and recipients' abilities to infer it." (924338, 24).

Importantly, the notions of "correspondence" or "mapping" seem to also fall out of the scope of Dahl's proposal, although it is interesting to consider sentences like "For the communication to be functional, the"meaning" imparted to a musical gesture by the performer must be recognized by the listener." (7789465, 36) as extensions of conversational repair considerations to a meta-level. Ultimately, Dahl's proposal accounts for an important part of the practice of ascribing meanings at the scale of organisms—namely, the role of understanding, highlighting the importance of semantic flexibility and dynamicity beyond the scope of relevance theory.

5.4.3 *The relevance of "information"*

An important element that is missing from the discussed views of semantics is the notion of "information", quite strongly present in the organism-level cluster (although not picked up by Word Sketch as collocated with the semantic terms considered). As mentioned above and discussed at length in chapter 4, "information" appears as a technical notion in results that are directly embedded in an information-theoretic contex:

> "In short, an information source produces a message, which is encoded by a transmitter into a signal that is suitable for transmission over a channel to a destination." (243943447, 51)

It also functions in a more relaxed, "everyday" meaning:

> "Information on the conversational partner's interest and understanding is also conveyed through gaze, allowing the speaker to adapt their message to enhance the listener's engagement and comprehension." (22133771, 39)

While there are some examples which equate "information" and "meaning", e.g.:

> "The signal is a stimulus that contains key information (meanings); the stimulus or emotion is modeled as a signal because it communicates a message or state and not because the signal is a wave function." (258461825, 81),

in most cases "information" is taken as a "vehicle" or "carrier" of sorts for the meanings. It is considered as independent from the semantic content:

> "The former communicates referential meaning by the means of intersubjectively understandable concepts (Bickerton, 2010), whereas the latter exchanges information which is at least more ambiguous in terms of its semantic content (Cross, 2005).", (2411028, 58),

but in many instances it is *about* something:

> "The ability to guide another individual's attention is a vital aspect of human communication, allowing us to convey information about what is important in the environment with ease." (222156320, 8)

> "In any case, it is undeniable that this form of nonverbal communication is extremely important for social cognition because it offers information to the rest of the community members about individuals' inner states, and it tends to emerge a few days after birth." (245297234, 32).

Both the relevance theory and the understanding-first approach, as originating from philosophy of language, presume certain higher-level cognitive abilities on the part of the communicating agents, and explain the emergence of semantic content of communication through those abilities. As we saw in the previous sections, the meanings in organism-level communication are shaped by speakers' and hearers' actions (and intentions behind those actions), their interdependence (as highlighted through the relevance of understanding), and embeddedness within a particular structure or correspondence between signals and their referents (whatever they may be). So where does this informational aboutness come into play?

In relevance theory, the concept of "informative intentions" is understood as aimed at changing the mental state of the recipient (Heintz and Scott-Phillips (2023); see also above in subsection 5.4.1). While "informative" is meant here in the more informal meaning (importantly, remember Scott-Phillips' rejection of informational accounts of communication that I have discussed extensively in section 1.2), it suggests that the mentions of information could be interpreted as addressing such internal changes in the recipient. Considering the appearance of plant signalling in the organism-level cluster, these might not be "cognitive" or "mental" states per se (though I'll ultimately advocate otherwise in chapter 6).

Nevertheless, it is quite common to assume that living systems, regardless of their mental faculties, do have some "information" about the environment. For instance, consider the cyberneticians' "Good Regulator Theorem" which offers a formal proof that for a system to effectively control its relation to environment, as organisms most certainly do (within some bounds), the system needs to be a model of the relevant states of environment (Conant and Ross Ashby 1970).

Alternatively, for those who prefer non-representational accounts, this feature is captured by the ecological psychologists' notions of "attunement" and "resonance" (Gibson 1986; Heft 2001; Raja 2018; see also Falandays, Yoshimi, et al. 2023). In that context, the mentions of information in the cluster related to organism-level meanings might be read as picking out the relevant intention on the part of the speaker and setting up the criteria of success for a communicative exchange.

In (a bit simplistic) informational lens, we can view the understanding emerging out of communication as an increase in the shared information between the speaker and the receiver. What is crucial—and what the informal uses of "information" overlook—is that this perspective requires a slightly different sense of "information" than a purely syntactic, Shannonian notion of information transfer provides (as explored in chapter 4). Consider the sentence:

> "To do so, we developed stimulus information representation (SIR), an information theoretic framework, to tease apart stimulus information that supports behavior from that which does not." (58625409, 10)

It contrasts the "stimulus information" which is likely standard correlational information (in the sense of natural signs, see subsection 4.2.4), with "information supporting (or not) the behaviour". This is also visible in modifiers and adjective predicates collocated with "information" in the results (as picked out by Word Sketch search of collocates for "information"): information can be "relevant", "useful", "contextual", "complete", or "incomplete"; "crucial", "necessary", "central", "fundamental", "important", "essential", "instrumental". All these notions only make sense in relation to a "semantic" (another important collocated modifier) notion of information. Aboutness introduces the criteria for usability or completeness: without the ability to normatively evaluate information, these properties are meaningless.

This, in a way, reinforces the broadly "encodingist" aspect of the discourse on meaning on the organism level. But it does so in a very convoluted way:

> "Communication occurs when the information changes the behavior of the receiver for the benefit of the signaler (Wilson, 1975, pp. 194-200)." (15924023, 490)

> "Information is exchanged via explicit messages utilizing language and shared representation, to bring understanding with each other." (216589085, 31)

> "If information's visual cues are perceived as more important than the information itself, then, now more than ever, 'the medium is the message'." (269507866, 74)

These are disparate examples, but what they all show is that the "encodingist" mapping—as I've already argued—is far from static. A message or a signal at organism level seems to be carrying always much more information than the receiver is supposed to and actually picks up. The "parts" (these can be physical dimensions or degrees of freedom) of the message that will actually impact the receiver depend on both communication partners' intentions and activities (and the broader conventions in which they are embedded)—and through that, become imbued with meaning. This ultimately bridges the Scott-Phillips' distinction of "informational" and "influence" accounts of communication (as I've already suggested in chapter 3): communication can involve both transfer or exchange of information (oftentimes, trivially) and be concerned with the influence of the sender on the receiver: the influence is enacted by (parts) of the information carried in at least some circumstances. Eventually, then, a meaning of any signal is some complex mixture of the sender's and receiver's intentions, inferences and the actual outcomes that the message achieves.

The only issue is that this account leaves us with the suspicious concept of INTENTION. In psychology, intentions often are explained in a "homuncular" form (e.g., "Behavioral intentions are instructions that people give to themselves to behave in certain ways", Triandis 1980, p. 203) and are studied through introspective methods, through "participants' responses that have the form, 'I intend to do X', 'I plan to do X', or 'I will do X' " (Sheeran 2002, p. 2). However, it is well-documented that introspective methods often fail to account for the actual psychological states (see Sheeran (2002) for the meta-analysis of the research on the gap between reported intention and actually performed actions) and do not have a clear epistemological foundation (for a review see Komorowska-Mach 2019). As such, even though I will argue in the conclusions (chapter 6) that the general goal-directedness INTENTION imputes, can be accounted for in a deflationary way (and considered widely present in the biological realm), and as such might not require the ascription of "higher-level" cognitive abilities, it is reasonable to consider whether intention can be omitted from this view. Indeed, this is especially the case since some readers of relevance theory argue for such a deflationary notion of intention, as referring to a "subpersonal, computational process" (Stegmann 2016).

The Floridi's data definition of information (introduced in subsection 1.3.1) may be useful in this regard. Floridi understands information as data (patterns of a lack of uniformity, distinguishable components in the world) that is well-formed (structured) and meaningful. For semantic information, he distinguishes between "instructive" information and "descriptive" or "factual" which adds the further requirement of veridicality (Floridi 2010, ch. 4). The emergence of such semantics, in his view, requires a particular cognitive architecture which implements an action-based semantics. The action-based semantics posits that "proto-meanings of the symbols generated by an [autonomous agent] are the internal states of that [autonomous agent], which in turn are directly correlated to the actions performed by the same [autonomous agent]" (Floridi 2011, p. 164; note that this refers indirectly to a much broader tradition of viewing semantics as grounded through action, e.g., Barsalou 1999; Grush 2004; which saw some uptake in robotics, e.g., Mannella and Tummolini 2023; Pezzulo et al. 2013).

The key point that Floridi makes, is that the particular version of action-based semantics he offers doesn't rely on any sort of perceptual representations and requires only the internal states of the system as "semantics-inducing resources". Further, these need not be guided in any particular fashion—the actions need not be purposeful to become semantic. The proposed cognitive architecture (Floridi 2011, pp. 166-176) specifies further requirements on a system that can achieve semantic contents from the action-based perspective (at the level required for solving the symbol grounding problem, Harnad (1990), see subsection 5.5.1 below; however Floridi's account fails in this regard, see Bielecka (2015)), including the requirement of a meta-level processing of object-level internal states.

This is where the potential advantage of Floridi's account lies: he postulates a mechanistic distinction between the internal states of a cognitive system constructing semantic content, which is "given directly every time an AM^2 [autonomous agent with the postulated cognitive architecture] interacts with the world" at the object level (Floridi 2011, p. 180), whereas the symbols (understood here as symbolic mental representations) are created when the object-level internal states are transduced (and abstracted) by the meta-level system. Ultimately, Floridi envisages that a complete account of semantics requires a network of such autonomous agents interacting with one another in a signalling game-like setup (see subsection 4.1.4. In particular, he proposes a very basic setup of a sender-receiver or reference game, where an agent sees an object to which another must react, and the sender needs to instruct the receiver what it sees, for both to receive a reward). In such circumstances, he argues, the Wittgensteinian problem of a "private language" can be escaped (Floridi 2011, p. 176; Wittgenstein 1968), as the immediate grounding in action becomes extended by this "social" system. In this, his points parallel a number of earlier positions (e.g., in the work of Luc Steels, reviewed in Steels 2015).

For the current purposes what is crucial is that the cognitive architecture implementing an action-based semantics is a particular specification of the "semantic context" for input data, more generally required for data to count as information (this characterization of Floridi's account is adopted by Dickins (2023), see also the discussion in subsection 1.3.1). For instance, for vervet monkeys we could argue that their communication system is semantic in Floridi's sense, as individual calls represent abstracted classes of events and are veridical (as well as well-formed), even if their content is likely instructional rather than factual.

Nevertheless, this view is difficult to apply to a description of real empirical cases of animal communication. Floridi's proposal is geared towards uncovering the very emergence of semantics and answering logical and epistemological puzzles, pertinent to philosophy of language. For many more complex organisms, such as vervet monkeys, the recurring example in this chapter, certain cognitive capacities can be already presupposed, which can trivialize the description in terms of Floridi's account. Monkeys can already be said to have semantic and symbolic representations, because of their particular cognitive architecture and an evolutionary history which, among others, led to the appearance of the particular communication system. While this history is not relevant in the same way as in teleological accounts (it is not *in virtue* of the selection history that signals have the meaning they have), the general trajectory of evolution and development needs to be included for a non-trivial application of Floridi's model to such more complex cases. Nevertheless, the signals—and the internal states which the action of signalling is associated with—have an action-based grounding (semantic content) in the previous instances of calling out and escaping in response to others' calls, and the internal states these actions were accompanied by.

Note that while these internal states are associated with actions, they have a distinctly non-intentional character. They explicitly lack any sort of motivational role (as required by the "zero" condition of no reliance on teleological concepts, accepted by Floridi as a necessary element of an accurate solution of the symbol grounding problem). If Floridi—or proponents of the closely related, alternative sensorimotor grounding approaches (e.g., Barsalou 1999), quite popular in some areas of psychology—are correct in this minimal understanding of semantics, the relevance theory's reliance on "intention" could be deflated, as semantics would not require any distinguished connection to *motivational* states, at the same time leaving much of the higher-level components of the theory intact. Hence, we can view the organism-level meanings as some complex mixture of the sender's and receiver's action-related internal states, inferences (logical or practical) and the actual (behavioural) outcomes that the message achieves.

5.5 CONSIDERING CELL-LEVEL MEANING

As discussed above, the empirical data identify a disconnect (if not a rift) between the scientific practices of ascribing meanings to signals in the organism-level and cell-level communication processes.

At the scale of cells, the imputations of semantics are much more indirect and subtle. In what follows, I'll assume that the sentences picked up by semantic search do refer to semantics simply in virtue of the methodology of the corpus study, even if the textual evidence is susceptible to different interpretations. As a support for the semantic interpretation we can take, for instance, a recurrent, though admittedly not overwhelmingly popular, appearance of modifiers indicating what the signal is about: Word Sketch across the whole results picks out "submission" and "stress" signals (with the former referring to the organism-level signalling, and the latter appearing both at the organism- and cell-level), and some further qualitative inspection of the sample picks out also "alarm signals". This suggests a clearly delineated class of signals that refer to the stress or danger that a cell might be in, providing some evidence for the claim that there's (ascription of) "aboutness" in cell-level BIOLOGICAL COMMUNICATION.

But the type of "aboutness" these modifiers indicate matches strongly the other, apparently semantically-oriented (as per the empirical results), and much more widespread practice of considering the signalling through what it achieves (as we also saw at the organism-level). Consider some of the examples (some already included in the box 5.2, here repeated for convenience):

1. "This inter-communication is important in the development and maintenance of the nociceptive neurons' hyperexcitable state, i.e., central sensitization." (267157380, 228; neuroscience);
2. "Cellular signalling is a complex process and involves cascades of enzymes that, in response to a specific signal, give rise to exact cellular responses." (268887219, 0; molecular biology);
3. "The T-cell developmental program is orchestrated by the inputs of multiple signaling networks." (256585371, 151; molecular biology);
4. "This work has been fueled by the claim that communication among the functionally specialized circuits is necessary for implementing complex, goal-directed behaviors." (258808097, 109; neuroscience, psychology);
5. "Bacterial cell-cell communication (quorum sensing, QS) represents a fundamental process crucial for biofilm formation, pathogenicity, and virulence allowing coordinated, concerted actions of bacteria depending on their cell density." (6479445, 0; microbiology);
6. "Under favorable environmental conditions, these signaling cascades control energy balance, cellular plasticity, and the mechanisms supporting homeostasis, growth, and reproduction." (6602278, 50; developmental biology, microbiology, molecular biology);
7. "In particular, one model for RasGEF action suggests that Ras signaling works as a complex network in which specific signals can be transduced

> by multiple RasGEFs, each of which can activate multiple different Ras proteins." (6133137, 143; ecology and evolution, microbiology, molecular biology)

Signals are considered as functioning within an input/output system (a functional recapitulation of the sender/receiver distinguishability, see chapter 3 and subsection 4.2.3), and the specific types of outputs they lead to seem to be what the signals are about. This is reminiscent of Kolchinsky and Wolpert's "causal necessity" account discussed in detail earlier—according to which the meaning of a signal is the information required by the receiver to maintain its existence. Interestingly enough, while these examples score highly on the "causal necessity" query, they receive even higher scores on other queries, in particular, the S-vector semantics one (despite these two queries having very close average similarity scores across the whole sample). Perhaps this is driven by the reference to "necessity" which is a strong and rare condition in biology (applicable perhaps only at the scale where physical descriptions of biological processes make sense—and as such, where those processes are explicable primarily through physical forces; this is the realm that Kolchinsky and Wolpert prominently feature in their account).[13] While at some points scientists are comfortable with claiming that communication "is necessary" for goal-directed behaviours (see example 4 above) or that it is a "fundamental process crucial" for concerted actions (example 5), most of the cases take the more cautious route of suggesting that signalling "contributes" to a variety of processes and that it "can" impact a variety of biological phenomena and properties.

Regardless of the degree of certainty, these various biological processes seem to form a more unified class of functions. Communication is seen as "orchestration", coordination or control of biological processes. Sometimes, instead of a functional description of this sort, the scientists opt for a more materiality-focused one, which centres a particular biochemical causal pathway (a distinction I will discuss below in subsection 5.5.2). This focus on control and constraints escapes the scope of linguistic relevance theory that served us well at the organism level. There is no talk of intention (even though we could, following proponents of basal

---

[13] For instance, Monod (1971) considers necessity in biological processes as the application of chemical and physical rules at the level of transcription and translation of DNA, though he highlights, as is now standard, the essential role of randomness and accident for evolution. Beyond evolution, Rosen (1989), in his careful rejection of physical reductionism—and of Monod's position—highlights that physical reductionism centeres the notion of physical necessity, which he couches in terms of "entailment" (using the category theoretic framework he has developed and finessed later on, see Rosen (1991)), which for Rosen doesn't have sufficient explanatory power with regard to biological processes. Instead, he opts for an alternative, often treated with suspicion, Aristotelian concept of final causality, and the notion of necessity/entailment that can be developed within that view. While I do not have space to cover his view in greater detail, it should be noted that his view is a naturalistic and mechanistic (though he has an idiosyncratic concept of a machine) adaptation of Aristotelian framework. Finally, the concept of necessity is closely related to that of a "law", which philosophy of biology has long abandoned (e.g., Brandon 1997).

cognition (e.g., Lyon et al. 2021), try to deflate this notion and apply it to cell-level phenomena—a direction I have already suggested within Floridi's framework and will briefly explore further in chapter 6), the "encodingist" mapping seems to be clearly defined (especially in the biochemical descriptions), as most strongly highlighted in example 2 above: a *specific* signal leads to *exact* cellular responses.

Nevertheless, it remains context-dependent (the signals' meanings will depend on what particular sender and receiver systems are involved in, and various "modifiers" might impact the signal along the way, changing its output and, in consequence, meaning). This seems to break down the notion of correspondence (at least in the "Fodorian" sense, the more deflationary accounts of correspondence, like the conception Miłkowski invokes (see Miłkowski 2023b), might remain intact): if biological signals indeed are about the processes they contribute to, their "aboutness" is embedded into their causal structure, with no "object" or "referent" outside that causal structure they would correspond to. But at the same time, this correspondence relation has troubled philosophers for quite a while (in the form of the problem of intentionality), with a variety of accounts, including the ones listed in the beginning of this chapter, aiming to capture how it is established. From this perspective, these empirical results resemble recent developments in those broader intentionality debates and may motivate abandoning a "mentalistic" notion of intentionality. I will now turn to those connections in order to sketch out a view of cell-level semantics.

5.5.1 *The symbol ungrounding problem*

In the context of mental representation, the problem of naturalizing intentionality has been prominently reframed as the "symbol grounding problem" by Stevan Harnad (1990). Harnad asks:

> "How can the semantic interpretation of a formal symbol system be made intrinsic to the system, rather than just parasitic on the meanings in our heads? How can the meanings of the meaningless symbol tokens, manipulated solely on the basis of their (arbitrary) shapes, be grounded in anything but other meaningless symbols?" (p. 335)

Tying this question directly to the (still pertinent) developments in cognitive science at the time, in particular the growth of connectionist models[14], Harnad is interested in understanding how formal computational systems[15] can "break

---

[14] Note the intriguing parallels between the early 1990s, with a resurgence of interest in connectionist model and the end of AI winter, and the early to mid 2020s, with the appearance and rapid growth of (societal) significance of transformer-based architectures and the generative artificial intelligence tools, and how questions about limitations of this technology come back (e.g., Bender et al. 2021; Birhane and McGann 2024).

[15] Connectionists—and the vast majority of cognitive scientists at the time—see the mechanisms of biological information processing as a computational process. This view has been

out" of a coherentist loop of symbols being grounded in other symbols, and successfully refer to something in the world. Despite clarifying the debate, his conceptualization of the problem did not ultimately give rise to a satisfying solution. For this reason, more recently, Terrence Deacon and Joanna Rączaszek-Leonardi proposed to reconsider this issue as a "symbol ungrounding problem" (e.g., Rączaszek-Leonardi et al. 2018; Rączaszek-Leonardi and Deacon 2018).

Building on Deacon's "Symbolic Species" model, embedded in Charles S. Peirce account of signs,[16] the "ungrounding" approach views symbols as inherently grounded at their emergence. Sign vehicles—in language or in the brain (though these authors do not agree with the dominant perspective on mental representation, see Deacon and Rączaszek-Leonardi 2019)—start off as grounded indexically and iconically (Peirce's first and second type of signs), through causal relations and structural similarity to the (behavioural) outputs or rewards they are grounded in, at the same time forming causal relationships with one another, which introduces an indexical grounding.[17] At this stage, there is already structure in the sign system, and the relations between the vehicles (what they call "syntactic" relationships) can be considered as similar to the relations between the structure of the outputs, i.e., these relations become iconically grounded (Peirce's second type of signs)[18], introducing complexity into the sign system. Deacon's model continues by highlighting that at this point the transformation of the referential function of the signs allows them to detach themselves from the immediate behavioural output: symbols (Peirce's third type of signs) appear as "ungrounded", with their reference being fully determined (at any particular

---

extensively debated and criticized (see Dreyfus 1992), but has offered important epistemic benefits over the years and remains the most prominent perspective on cognitive processes. Importantly, some of the criticisms that have accrued over the years can be solved by the alternative view of Rosen (1991), as e.g., Tomasz Korbak explores (2023).

[16] While the precise views of Peirce on signs (e.g., Peirce 1985) have changed over the course of his life, he saw semiosis as a triadic process throughout his work (my presentation here follows Atkin 2023). He distinguishes three central components of a sign: its signifying components, variously called the "sign" or "representamen", the "object" which the sign represents, with the object determining the sign, and the "interpretant" which he has variously defined in terms of understanding or grasping the representamen-object relation, or a translation or development of the sign. His early account postulated semiosis as an infinite process, since the interpretant was considered as a sign itself—again to be divided into the three components; later he modified his views to avoid this infinite regress. Later accounts offered additional categorizations, e.g., distinguishing types of objects and of interpretants, which aimed to capture different stages in semiosis seen as a dynamic process. Throughout his work, he also distinguished three types of signs: indices, icons, and symbols, which differed by the relation between the representamen and the object—as discussed in the main text.

[17] Rączaszek-Leonardi (personal communication) argues that the processes of iconic and indexical grounding happen concurrently, without a clear hierarchy, and both are necessary to ground symbols.

[18] Note that while motivated by different assumptions, the view at this point resembles in important regards Shea's varitel semantics and Miłkowski's correspondence theory, discussed in the beginning of this chapter: the possible iconicity of sign-sign relations to relations between their objects can be understood through the notion of structural correspondence, providing an "exploitable relationship" for representation.

moment) by the structure of the sign system and its context. But the historical trajectory of their development—across both the phylogenetic and ontogenetic timescales—explains away the "mystery" of their detachment.

Deacon's work focuses on the evolutionary timescale (e.g., Deacon 1998, 2017), whereas Rączaszek-Leonardi and her collaborators' studies elucidate how this process operates on the developmental timescale (e.g., Rączaszek-Leonardi 2016; Rączaszek-Leonardi et al. 2018; Nomikou et al. 2016; Nomikou et al. 2017). To illustrate the ungrounding, let's consider the examples of mother-infant interactions Rączaszek-Leonardi analyzes. She works on the data collected by Katharina Rohlfing and Iris Nomikou, a longitudinal (with visits at 3, 6, and 9-months of child's age), multilingual and multimodal (video and sound recordings, transcribed and coded) corpus of mother-infant interactions outside of laboratory conditions, documenting structured activities such as diaper changing. Through microanalyses of the behaviour of participants (e.g., Rączaszek-Leonardi, Nomikou, and Rohlfing 2013) the authors were able to retrace the temporal dynamic of the development verbal interactions. The overall findings suggest that already with preverbal infants, mothers introduce a dialogical structure of interaction, in which they treat the behaviours of the children as responses to their verbal or non-verbal cues, as well as the causes of their own behaviour, embedding their activities in a social structure.

Consider the example of a mother putting shoes onto the feet of a 3-month old boy, after changing the diaper (as described in Rączaszek-Leonardi, Nomikou, and Rohlfing 2013, pp. 7-9). The interaction begins by the mother announcing her action, showing the shoe to the baby, pausing, and gazing at the baby's feet—she resumes her action only after the boy has kicked his feet. As one of the feet was raised higher, the mother interprets this as the child making a decision with regard to which foot should go first. She comments on it, clarifying if she understood correctly, and proceeds to put the shoe on, after which the same structure of actions is repeated for the other shoe.

Rączaszek-Leonardi, Nomikou, and Rohlfing (2013) point out two central features in this example (which recur across other interactions identified in their corpus): first, while the baby is too young to understand mother's utterances, his "behaviour is treated as originating from his own agency, relevant to the situation. The infant's movements are interpreted as proposals or decisions that influence the outcome of the interaction." (p. 9). In this way, already at this stage of development, the child is learning to "identify goal-directed actions and intentional agents because [he] witness[es] agency within communicative actions" (p. 9). Crucially, the mother provides a time window for infant's actions, introducing a temporal structure that persists through a vast majority of (polite) interpersonal interactions, and the active participation of the boy in the interaction "may also serve as an additional ostensive cue that educates the infant in becoming an intentional agent" (p. 9). Second, as the mother repeats the action for the second shoe in a closely similar manner, regularities of the interaction emerge and serve to constrain future interactions of the same sort: a "frame" is established, within which certain expectations become clear (e.g., of the child raising one of the feet rather than kicking them randomly), eventually leading to conventionalization of the behaviour.

From the perspective of ungrounding, the crucial element of this example is how an initially random action of the baby boy, the kicking of the legs, is interpreted by the mother as an intentional decision on his part. Indexicality, a causal link between infant's and mother's behaviour, is established, and the repetition of the structure allows for the establishment of iconicity: similar actions between recurring interactions become related, eventually allowing the child to e.g., raise its foot outside of the diapering circumstances, as a way of communicating its intention to have a shoe put on. Further, the process occurs in parallel also at the level of mothers' narration, which is progressively differentiated from other modalities through its own timing and repetition. Thus, the child is exposed to the fact that structured language can also act intentionally, and that interaction partners take it as an intentional action in the (social) world.

Ungrounding cell signalling

Going back to signalling at a cellular scale, the "ungrounding" perspective offers significant epistemic benefits. In the case of biological communication at the cellular scale, there is no formal systems for which we may consider a problem with "grounding". This may lead some theoreticians more strongly tied to the assumptions of the "encodingist" stance to reject the claim that these processes constitute a case of semantic signals, and argue that the whole "communication talk" at cell level is strictly metaphorical and ultimately should be abandoned (a possible couterargument to my view, which I considered in detail in chapter 4). On the other hand, this gives us the opportunity to consider signals that are inherently grounded (in virtue of their causal structure), in a similar manner that the infants' (presymbolic) actions (that Rączaszek-Leonardi and her collaborators consider) are.

To add detail to this point, consider again the case study of metabolic integration of *B. subtilis* biofilm orchestrated by potassium signals (discussed in detail in subsection 3.5.1). Prindle et al. (2015) frame the potassium signalling in the input/output, functional terms (that, as I argued above, can be interpreted through a semantic lens), and they link the potassium signals to metabolic stress:

> "The initial trigger for potassium release is metabolic stress caused by glutamate limitation. External potassium depolarizes neighboring cells, producing further nitrogen limitation by limiting glutamate uptake, and thus produces further metabolic stress." (Prindle et al. 2015, caption for Fig. 3)

The study highlights that the signal is actively propagated, and that its success requires the presence of the YugO ion channel. In both wild-type and the yugO deletion strain, the scientists observe that upon a change in extracellular potassium concentration, there's a membrane depolarization. However, in the wild-type but not the mutant strain the initial depolarization is followed by hyperpolarization and an increase in extracellular potassium. The influx of potassium ions enabled by the presence of the YugO channel (rather than the depolarization) is recognized as a signal and propagated, and as its function persists across the propagation—through the reduction of metabolic rate of the recipient—we can consider it as preserving its 'meaning' (I will use scare quotes to refer to 'meanings' to highlight that the substantivist, encodingist assumptions that the term carries might not be accurate for the current case).

This 'meaning' is inherently grounded in the causal structure of the process: the signal results from metabolic stress and leads to metabolic stress.[19] This gives us a very elegant identity mapping from the signal to its 'meaning', but the presence of *active* propagation of the potassium wave suggests a more complicated picture: one that ultimately leads Prindle and colleagues to conclude that those signals communicate the metabolic state of the cell, a slightly more abstract "referent".

In Peircean categories, the potassium signal is an index, but the emerging propagation could be seen as introducing a "secondary" indexical grounding, by connecting the signals to one another. This is a first step of an ungrounding process, as captured by Deacon's framework. And while *B. subtilis* biofilm remains at an indexical level, with the signals inherently grounded in the causal structure of the process, we see semantic properties appear—through the functional-causal intertwinement of signals and cell activities, and the way that signals act as constraints on the behaviours of individual bacteria within the biofilm.

Constraints and structure

This is undeniably a very different way of understanding meanings than the one that emerged from the organism-level examples. Nevertheless there are several important points of contact: the role of convention (in the form of intra-systemic structure), or the focus on functional outcomes (visible in Dahl's "understanding-first" proposal). I will return to these similarities shortly, but let's explore the differences first.

At the organism-level, the meaning of a particular signal is established by the representational internal state of the sender and the way it can be encoded in the signal and identified by the receiver (or impact the receiver's internal state). No such cognitive resources can be indisputably claimed at the cell level (although, as I've mentioned before, I will explore an alternative view motivated by basal cognition further down the line). Instead, as we've just seen, the semantics of communicative processes are established through the causal structure of the

---

[19] 19. The coordination of the response could be also accounted for in terms of a signalling game: the payoff—survival—is awarded at the level of the colony, hence the sharing of information allows the bacteria that are not (yet) nutrient-deficient to learn about the looming danger (as even if they'd never lose access to ammonium, the death of their kin leaves the biofilm vulnerable).

process, and the active role that individual cells play, particularly as cashed out through the notion of constraints. This is precisely the view that Howard H. Pattee has advocated. Importantly, as his work happens in parallel with the major developments in understanding genetics and evolution which have shaped these disciplines for several decades (e.g., Dawkins [1976] 2006), the majority of his examples considers the role of DNA and the genetic code in the development of cells, even if his conclusions are broader than this.

Pattee's goal is to explain the "symbol-matter problem", his way of restating the problem of naturalizing intentionality, that has been returning throughout this chapter. Pattee is an evolutionary gradualist with regard to mental capacities: he accepts that the familiar mental predicates of "intention", "belief", "thought", or "meaning", must have had simpler biological predecessors. And as such, these predecessors offer unique model systems for studying the properties they share with those mental predicates, without the complexities of human psychology. He traces the simplest such systems through the course of evolution to individual cells and the mechanisms they have available, arriving ultimately at a systemic view in which meaningful messages emerge "only in the context of a larger system of physical constraints called a 'language' " (Pattee [1969] 2012, p. 61).

As he summarizes, "a language must be a closed set of structures, which we call symbols ("signs" in the Peircean parlance I've accepted throughout; Pattee, personal communication), some rules for joining and transforming these symbols which we call the grammar, as well as a set of rules for interpreting the meaning of such a collection of these symbols" (Pattee [1972] 2012, p. 81). Crucially, in his view, these systems came about from "a primitive communication network which could be called the primeval ecosystem", e.g., a "primitive geochemical matter cycles in which matter is catalytically shunted through cell-like structures" in a spontaneous, uncontrolled manner (Pattee [1969] 2012, p. 60). In such systems meaning emerges through a combination of syntactic structure and semiotic closure.

Syntactic structure operates by introducing (hierarchical) constraints: one of the key concepts in Pattee's broader theoretical framework. Considering life as a non-ergodic, non-stationary process—i.e., one that is causally dependent on past states, what Pattee calls "records"—he notices that life evades standard physical theory precisely as physics lack tools to describe the continued generation of hierarchical, coordinated constraints. These constraints "write and read records" (Pattee [1973b] 2012, p. 70), i.e., enable the causal efficacy of the biological systems' historical trajectory, and as such they have accompanied life from its very inception, "from the rules of the genetic code to the rules of the languages of man" (Pattee [1973b] 2012, p. 70).

Throughout his work, Pattee develops the notion of constraint, tying it to the problem of measurement in quantum physics (e.g., Pattee [1972] 2012, [1973c] 2012) and embedding it in the formal framework of hierarchy theory (e.g., Pattee 1973a, 2017; there are important parallels to the work of Robert Rosen and his use of category theory, a similarity that Pattee also notes at some points).

This leads him to distinguish and characterize "control constraints" as a separate class of hierarchical dependencies: "developmental controls in cells may be executed by 'ordinary' molecules to which we give titles, such as activator, repressor, or hormone, but the control value of these molecules is not an inherent chemical property; it is a complex relation established by a collective hierarchical organization requiring the whole organism" (Pattee [1973c] 2012, p. 94).

Syntactic constraints are precisely such control constraints: they introduce a complex structuring of possible interactions, channeling the semiotic interactions into particular causal pathways. Pattee considers "molecular information processing", the process of reading, transcribing and translating DNA, precisely in those terms. His central point rests on the arbitrariness of the genetic code: "there appears to be no physical or logical reason why the same phenotype could not in principle be produced by gene strings in a very large number of different genetic languages" (Pattee [1982] 2012, p. 171). Hence, if the "meaning" of genetic code is the phenotype that is produced on its basis (with all the caveats regarding genetic information discussed in subsection 1.3.1), this meaning is possible only through the complex role of DNA as a control within a broader system of interacting biochemical components, rather than through some inherent structural properties of the nucleotides.

At the same time, Pattee's framework requires the contraints to be *replicable*. This notion is discussed in depth by Joanna Rączaszek-Leonardi (2012), who discusses it in terms of transmittability, which is the necessary condition for a physical structure to be selected (in the course of evolution) as a functional constraint. Such structures require a particular physical makeup, which enables them to act as controls, functionally constraining a dynamic process, at the same time allowing them to be copied as a purely physical system—in a relatively energetically cheap way. DNA is a standard example here, as it acts as a constraint on the processes of translation and transcription, as well as can be split and effectively copied. But replicability introduces important conditions when considering natural languages as well, as for a verbal utterance to be replicable it should follow the principles of least production effort and greatest perceptual distinctness (Rączaszek-Leonardi 2012, p. 312, see also the role of repetition in Rączaszek-Leonardi's analysis of symbol ungrounding in mother-infant interactions above). This underscores the importance of particular material implementations of the communicative systems.

For meaning to emerge, syntactic structure must be further coupled to the principle of semiotic closure. Semiotic closure (which Pattee initially called "semantic closure", see Pattee 2012) boils down to the introduction of perspectivality of the sign meaning: the referent is established as the sign is "interpreted", but the processes of "interpretation" itself depend on the semiotic characteristics of the sign. Effectively, the sign functions in a self-referential loop (see Pattee [1982] 2012).

A crucial element of semiotic closure is that this self-referentiality depends on the physical embodiment of the signs and the interdependence between those embodiments is what drives self-referentiality. The controlling molecule is part of the process being controlled, and its ability to constrain that process depends precisely on belonging to the process. Pattee centers the case of nucleotides in a DNA sequence being "directly recognized by translation molecules (tRNAs and synthetases)" (Pattee [1982] 2012, p. 172), but we can also think about this in terms of the structural matching between a biochemical messenger and the relevant receptor protein on a cell membrane. Only within such semiotically closed process, the syntactic structures of biological signs can be endowed with meaning (and, effectively, functional).[20]

To come back to the potassium case—in Pattee's view the direct causal impact of the ion signal is considered through the lens of "semiotic closure". The ion signal is a self-referential process, that is made possible because of the particular syntactic structures present (ion channels, the charge of glutamate and ammonium). It acts as a complex structural constraint because of the physical dependence of the rate of uptake of nutrients on membrane polarity and the voltage-gating mechanisms in the ion channels. Within that system, the signal can be received and interpreted, i.e., functional. As the process is realized precisely through the physical embodiments (electrical charge, particular molecule size of potassium) of the signal, it actually effects the expected outputs.

5.5.2 *Constraints, interpretants, and cell-level semantics*

To go back to corpus data: biological communication and signalling are primarily seen as a way to control biological systems, in particular as a coordinating and integrating activity, for example when individual bacteria and their engagement in quorum sensing processes enable and guide the processes of biofilm formation, or when it connects specialized circuits (e.g., in the brain or in development) enabling them to work together to achieve complex, goal-directed outcomes. These processes of control and integration with their—often explicitly recognized—goal-directedness seem to be what individual messages *refer to*, what they *mean*. And while such openly semantic notions are rarely (though not never) used in reference to cell-level communication phenomena, as the empirical results indicate, indirect ascriptions of meaning happen precisely through the identification of the functional outcomes of individual communicative exchanges.

---

[20] The concept of "semiotic closure" is embedded in a broader web of various forms of closure that are explored by a set of theoretically related positions in philosophy of biology and philosophy of mind: Robert Rosen's idea of "closure to efficient causation" (Rosen 1991), enactivists' "operational closure" characteristic of autopoietic systems (Maturana and Varela 1980; Varela, Thompson, and Rosch 1991), and the closely related idea of "organizational closure" considered by the proponents of the organizational approach to biology (e.g., A. Moreno and Mossio 2015; Mossio, Bich, and A. Moreno 2013), discussed in section 1.2, or the concept of "recursive self-maintenance" in Mark Bickhard's interactivism (Bickhard 2009). All these concepts underscore the relevance of self-production as apparently unique to biological systems, but differ in how the physical, material makeup (versus structural and organizational properties) of the system come into play.

The ecological framework of meaning emerging from the work of Deacon and Pattee (and several other authors that I did not have enough space to consider in detail, for instance Susan Oyama, e.g., Oyama (2000), or the more recent enactivist work led by Hanne de Jaegher and Ezequiel di Paolo, e.g., De Jaegher and Di Paolo (2007), Di Paolo, Buhrmann, and Barandiaran (2017), and Di Paolo, Cuffari, and De Jaegher (2018)) offers us conceptual tools to capture those characteristics of cell-level semantics. The meaning of individual biological signals at this scale is an effect of the complex structures of biochemical interactions—factual or possible—in the operation of the cell, and the function that those signals play, crucially, for the system itself.

Both of those aspects can be modelled in terms of constraints, and exhibit irreducible systemicity, in which any individual message depends on the presence of a larger system of structures within which it acts. A signal integrates and controls a system by constraining possible future trajectories and nudging the system—as a set of biochemical reactions—in a particular chemical and energetic direction. This is done purely materially (or physically and chemically), but the success of the signal depends on the presence of that systemic structure, on the signal being able to interact with other components of the system. And within such a structure, a formal description in terms of "information" is perhaps possible, but any such information will be (at this level) inherently meaningful, grounded—"a difference that makes a difference" (Bateson [1972] 1987, p. 460) for the system in question.[21]

The Peircean triad, his characterization of the process of "semiosis" (see the discussion above in subsection 5.5.1, as well as in Pattee (2021); Pattee (2023)), helps us capture the systemic character of signalling in different terms. Consider what happens in biological signalling when the potassium signals in the biofilm integrate the metabolic state of individual bacteria, or when the Notch receptors bind a ligand from a neighbouring cell and the receiving cell modifies its gene expression to switch to a particular cell type in the course of further development. The sign—potassium or the ligand—is interpreted in this case through the particular biochemical sequence of events it enacts, and in this way it is "understood", connected with its object—the output of that sequence, e.g.:

> "Specifically, the functions of signaling networks are based on the spatiotemporal dynamics that emerge from extensive protein groups and interactions collectively, rather than the simple presence or absence of individual proteins or interactions." (234743417, 31)

---

[21] Bateson originally defines an elementary unit of information in this way, trying to distinguish information from its physical vehicles, as the the analysis by Sloman (2018) highlights. Sloman is critical of the success of this account. In his (1979), Bateson makes a slightly more nuanced point that a difference "becomes information by making a difference" (p. 68), highlighting that information is a relational, necessarily epistemically accessible property. MacKay (1969) makes a closely related point, though in slightly different terms, in highlighting that "information [is always] *about* something" (p. 158).

But since the interpretant and the object are intricately tied to the sign's physical embodiment (in a way that it doesn't make sense to distinguish the sign from its vehicle; but note the point about material discontinuity emerging from the analysis in subsection 3.5.1), a self-referential loop is created. The sign is its own index, and it is a sign only when this indexical—so causal—connection is possible. Crucially, this is possible only inasmuch as we consider the recipient cell as a dynamic, active (or even agential) process: the cell is not a passive input/output function that awaits the signal and acts on its reception and then halts (in a way all too well known from various artificial systems). Instead, the cell is engaged in a set of activities that create the systemic structure within which the signal is meaningful—the Peircean interpretant. And the signal's meaning consists in how the (chemical, energetic) trajectory of that systemic structure changes in response to that signal.

To restate it again: the meaning of a cell-level biological signal is the control it enacts over the recipient system, but for that control to be possible, the recipient system must be an active system in its own regard, sensitive to the signal. This makes biological meaning at the cell level normative: if the signal is not received (e.g., in the mutant case in Prindle et al. (2015) experiments, when the bacteria are genetically modified so that the ion channel is lacking or not sensitive to the extracellular potassium concentration), the communication fails and it turns out that the signal was meaningless (for that particular system; a mechanism of this sort may be a factor in carcinogenesis, see Carvalho (2021)).

Further, we can also consider instances of falsity or dishonesty (although some additional conceptual tools of a particular theory of biological function are needed in that case), when a particular signalling system is being "hijacked", e.g., in the case of plant galls induced by various parasitic insects (among others) in which case gall formation (unusual cell differentiation and growth) is induced and controlled through continuous signalling from a larva of that insect which develops inside the gall (e.g., L. J. Harper et al. 2004; Raman 2011). The signal is "false" or "dishonest" in that case as there is a mismatch between its function *for the system*, and the actual outcomes of a given biological process (as mentioned, a particular interpretation of that distinction depends on the notion of "biological function" we accept).

Effectively, the account of meaning in terms of constraints or control provides us the necessary conceptual tools to account for the cell-level semantics of biological communication. However, an important question remains of what is the relevant "level" of the referent.

Fine-grained functionalism[22]

At various points in the foregoing discussion, I have switched between considering the meaning of cell-level biological signals either in terms of their functional outputs, or the particular biochemical pathways they control. This raises two questions: one is the question about a relevant notion of function that I have already briefly discussed in the first part of this chapter and that I will return to when synthesizing the empirical results. The second is the question of the relevant level of analysis, to which I turn now.

The selection of an appropriate level of analysis is crucial for how neuroscientists identify and individuate functions they aim to account for. This process—and more broadly, the process of identifying and individuating the targets of our models in science—tends to be overlooked by methodological prescriptions scientists develop themselves. Philosophers are more sensitive here, and they realize the relevance of this problem for several important questions. In philosophy of biology, this is visible in the evaluations of the concept of gene "homology" (e.g., DiFrisco 2023; DiFrisco, Love, and Wagner 2020, 2023), and more generally, some feminist philosophers of science and proponents of social constructivism have explored the impact of social factors on the target-identification procedures in science (e.g., Haraway 1992). In the context of cognitive science, these debates recur under the guise of establishing a "cognitive ontology", a taxonomy of mental processes and functions (e.g., Craver 2009; Khalidi 2022; Krickel and Goddu 2025).[23]

The trouble with standard taxonomies of cognitive categories is that they remained largely unchallenged since the heyday of theorizing about cognitive function. While a systematic treatment of their history is lacking, the "default"[24] set of cognitive categories (e.g., memory, learning, perception, inference, decision making, etc.) has varied origins, being partially inspired by folk psychology, and partially by historical speculative or introspective considerations regarding human cognition (see Buzsáki 2019, pp. 6-7; Vanderwolf 2007, pp. 1-4). Most importantly, it has gradually accrued over decades or centuries of research (see e.g., Anderson 2015; Vanderwolf 2007, pp. 1-4), with some researchers even suggesting their origin reaches back to Plato's division of the soul into three parts (see McCaffrey and Wright 2022).

---

[22] Parts of this section pull on the views about the epistemic benefits of fine-grained functionalism with regard to the material basis of cognitive function that we have developed and explored with David Harrison and Urtė Laukaitytė in Harrison, Rorot, and Laukaityte (2022).

[23] I'm thankful to Jakub R. Matyja for suggesting I tie the issues discussed here to the cognitive ontology debates.

[24] I'm opting for the term "default" (in a limited, Western—European and Anglophone—sense), since while these categories have come to be accepted as our folk psychological taxonomy, their varied origins suggest that calling them "folk psychological" full stop might be misleading—consider the concept of "motivation" that Anderson (2015) discusses, that have originated from scientific research only to then be broadly accepted as part of our "popular" knowledge about the mind. Instead, "default" accentuates their obvious, widespread use, without presupposing their origins.

Philosophers have realized these limitations of the default categories, and the "cognitive ontology" literature aims to rethink practices involved in the categorization of mental processes (e.g., Anderson 2015; Craver 2009; Khalidi 2022; Krickel and Goddu 2025). One success story comes from taxonomies of memory, that have been evolving in response to empirical results (see e.g., Allen 2017; Khalidi 2017; Werning and Cheng 2017). If there's any inductive conclusion to draw from this example, it is that these "default" categories do not clearly track real properties or natural kinds (see again Khalidi 2022; Craver 2009, also highlights the conventional elements in mechanistic delineation of natural kinds), and need to be reconsidered in the light of empirical results from neuroscience and cognitive psychology.

Beate Krickel and Mariel Goddu (2025, p. 2; see also "splitting" and "lumping" highlighted by Craver 2009) identify two main approaches to identifying cognitive categories in the scientific and philosophical literature. Radical or bottom-up views reject extant categorizations of function and seek to construct alternatives, usually beginning with neural categories, e.g., anatomically identified brain regions. On the other hand, incremental or top-down approaches accept the categorizations currently employed as a starting point and instead seek to evaluate their adequacy and usefulness, and improve upon them. Such is the approach accepted by Krickel and Goddu.

However, a limitation of incremental approaches is that with the historical categories, they inherit a particular level of granularity of the description of cognitive function. A related problem can trouble the "radical" approaches, since the anatomical practices of identifying brain regions are historically intertwined with the historical atlases of functions. While this does not defeat any of these approaches immediately, it is an important concern in more extensive comparative projects, like the current one, where considerations of scale are particularly pertinent.

Some theoreticians have problematized this. Peter Godfrey-Smith (2016), in arguing for the role of small scale material dynamics in cognitive systems, has provided the tools to begin to capture the role of scales for the taxonomies of cognitive function.[25] He distinguishes the dominant "coarse-grained func-

---

[25] Godfrey-Smith positions his debate in the context of the functionalist theories of mind (J. Levin 2021). Functionalism argues that mental states are defined through their function, independently from their physical implementation (known as the "multiple realization", a topic of intense philosophical debates, e.g., Polger and Shapiro (2016)). While functionalism with regard to the mental capacities is one of the most widely shared views in philosophy (see Chirimuuta 2018), there are important criticisms. For instance, in a vein related to Godfrey-Smith, Mazviita Chirimuuta (2018; 2024) argues for taking functionalism as a purely methodological, instrumental strategy, which is interpreted metaphysically in an unwarranted way.

A standard alternative to functionalism is the type-identity theory, advocated for by e.g., Polger (2004). Type identity theorists argue that (at least some) types of mental states are identical to (at least some) types of physical states, which has slightly stronger reductivist implications than the functionalist approach. Importantly, the historical definition (see e.g., Schneider, n.d.) has drawn the identity between mental and brain states, but more recent versions commonly accept that cognitive processes are not skull-bound (e.g., Miłkowski et al. 2018), and expand the class of tionalism" from the "fine-grained" alternative he advocates for. The standard functionalist account tends to consider processes such as "learning", "perception", "decision making" as its targets. But these cognitive categories are defined at a particular (spatiotemporal) scale: they emerge as a particular coarse-graining of a host of finer operations of the underlying systems. Such coarse-grainings are unavoidable in scientific practice. The scale is closely tied to particular experimental techniques, and it impacts how "real" are the properties or processes it identifies (see Flack 2017).[26] For instance, in the context of neuroscience, single-cell recordings, EEG, and BOLD [blood-oxygenation-level-dependent] imaging are all coarse-grainings, but our current best theory of neural activity suggests that single-cell and EEG recordings pick out causally efficacious properties (single cell more so than EEG), as they allow recording of cells' electrical activities. BOLD, instead, targets purely epiphenomenal characteristics of underlying, finer-grained biological processes, as it measures activity through the changes in blood oxygenation—which is not directly causal for neural processes, but rather a by-product of neurons' electrical activities.

This motivates the emphasis on finer-grained properties (compared to the default categories). For instance, in his discussion of how the idea emerged that bacteria possibly have memory, Colin Allen rejects some philosophers' insistence on distinguishing "genuine" cognitive capacities from "simpler" mechanisms that—in their view—bacteria engage in (cf. Lyon 2020). Instead he argues for a form of pluralism regarding cognitive taxonomies, but he does so precisely through the fine-grained properties of the putative categories. For memory Allen notes:

> "Many scientists (and some philosophers) are, however, beginning to understand that the molecular binding mechanisms regulating communicative and adaptive capacities of bacteria are deeply homologous, structurally and functionally, to the neuronal NMDA-receptors central to synaptic long term potentiation in learning and memory." (Allen 2017, p. 5; see also Stotz and Allen 2012; see also the considerations regarding cognitive homology in Krickel and Goddu 2025)

In his view, neural and bacterial memory are the same scientific kind in virtue of these mechanistic parallels in how they are implemented. More generally, he notes that (especially in biological sciences) a clear delineations and exact categories are often lacking across sciences and researchers rely more on "common questions and techniques": "Biologists don't need a definition of 'fish' that

---

possible realizations of mental states. Nevertheless, the considerations regarding coarse-graining of mental categories are relevant for both funtionalism and type-identity theories.

[26] "Coarse-graining" is an important element of the processes of idealization and abstraction that permeate science (among others, see L. Nowak 1980; N. Cartwright 1983; Chirimuuta 2024; L. Nowak 1972; Potochnik 2017; Weisberg 2007; 2013; Wimsatt 2007, see also the discussion in subsection 4.2.3), and the full picture of its role would require a more thorough discussion of those debates than I can get into here.

abstracts from all […] variety. Likewise, I maintain, cognitive scientists need neither an abstract definition of 'cognition' nor a theoretically pure conception of 'cognitive system' " (Allen 2017, p. 5). We could go back to the notion of "patchwork concepts" to capture his idea of pluralism about cognition, but what is particularly relevant here is the insistence on techniques and fine-grained properties they target. Allen rejects the idea that coarse-grained cognitive categories could form natural kinds (though see the approach suggested by Khalidi 2022), but at the same time accepts the finer-grained commonalities between various systems and strives to elucidate them (a point that we explore further in Harrison, Rorot, and Laukaityte 2022).

Coarse-graining cell-level meanings

In the context of cell-level semantics, fine-grained functionalism offers an approachg to the problem of distinguishing the appropriate "level" of the referent to consider. Above, I have suggested that cell-level biological signals are about the processes they control. However, these processes can be characterized through many different coarse-grainings. In the case of potassium signals in biofilms that I've been returning to throughout, we can claim that the signals are about the metabolic states of their senders, the metabolic state of the colony (that some bacteria perceive directly, and others need to be "told" about), or about the particular pathway of processes they trigger (membrane depolarization, associated with the reduction of uptake and retention of the positively charged ammonium molecules; and the following hyperpolarization triggered by the YugO ion channel). In the case of Notch signals, we can consider them to be about the "direction" of differentiation (e.g., proliferation vs. differentiation in neurogenesis) in a given area of the developing organism or about the particular gene expression change that the intracellular domain of the Notch protein triggers upon release (in both cases the meaning is likely context-dependent in various ways, e.g., the cells respond differently depending on spatial and temporal cues, by either stopping or promoting differentiation, see Chambers et al. 2001).

In a sense, this tracks the problem of distinguishing proximal and distal content of perception (see Dretske 1986; Davidson 2005). However, the direct issue at play here has to do with the scientific practice rather than metaphysics. As mentioned above, our ascriptions of meaning will depend on the particular notion of function we accept, but as we saw in the beginning of this chapter, a majority of those philosophical naturalizations do not provide us with conceptual tools for individuating functions. At the same time, Pattee's work on constraints underscores that they're meaningful only within a particular set of structures, with their material self-referentiality.

What this suggests is that those meanings will depend on how we carve out Pattee's "languages", sets of structures, a practice that in turn depends heavily on the particular aims and assumptions that researchers bring to the table (their perspectives, in the sense of Massimi 2022). But the view of constraints-based semantics advanced here requires these ascriptions to include finer-grained material properties. Potassium signals in biofilms may be about the metabolic states, but they are about them in a particular, material way. Notch signalling may be about the future trajectory of a cell, but this aboutness is structured through the material-causal pathway that leads to that future. This allows us to include the various "modifiers" or context-dependent effects that recur through scientific descriptions of biological semantics directly into the proposed view: by considering them through their material substrates.

This foregrounding of the materiality of signalling processes for their meaning underscores the importance of the material discontinuity necessary for a particular causal process to be considered an instance of communication (as I've argued in chapter 3). This is the step that we could identify as a Peircean interpretant, and if we hope to understand the meaning of a particular signal *for the system* in question, this is precisely the focal point for the analysis. But at the same time, the material self-referentiality emerging from Pattee's account ensures that the biological signals physically ground their meanings.

5.6 SEMANTICS OF BIOLOGICAL COMMUNICATION

While I have argued throughout this dissertation that there's a unique concept of communication across the levels of biological organization, the picture emerging from this chapter complicates this view. Empirical results clearly indicate that practices of ascribing semantics differ between what can be termed the "organism-level" and the "cell-level".

When we consider the interaction of multicellular organisms, a topic cluster that groups together human language and non-linguistic forms of communication across various taxa, a relevance-theoretic approach seems to capture the practices of ascribing meaning well enough. In that context, researchers consider signs as referring to "something" (objects, properties, or states, external with regard to the language) in structured, if complex and dynamic, ways. These mappings are enacted by the sender, embedded in their cognitive states and prior conventions, in ways that can, and must for the communication to actually be meaningful, be actively picked out by the receiving organism (in a way that enables understanding). The "aboutness" is ensured by the cognitive states (including intentions and inferences) of communication participants within the set of (mutually accepted) conventions.

When we look at the cell level, namely at interactions of unicellular organisms or individual cells within an organism, a constraint-based accounts of semantics is more appropriate. The signals are meaningful because of the way they physically interact with the ongoing processes in the recipient system. This is not an intrinsic, but rather relational property: the meanings are not inherent in the signals, but rather *arise* when the signs function within a larger set of material structures. The meanings consist in how the signals control—change, constrain—the development of those recipient processes. But ultimately, if the signals have any aboutness at all, this aboutness is grounded in the materiality of the signal within that set of structures.

As I have suggested briefly above, there are points of contact between those wildly different accounts: the emphasis on the active reception of the signal, on its embeddedness within a larger structure or convention, the dynamic view of the relation of the message to its meaning. These are not superficial or trivial commonalities; in fact, they align closely with the growing body of work on non-representational, ecological theory of language—despite the theoretical divergences within that work (e.g., Di Paolo, Cuffari, and De Jaegher 2018; Rączaszek-Leonardi 2023; Rączaszek-Leonardi et al. 2022; Rączaszek-Leonardi and Zubek 2023; Główka, Zubek, and Rączaszek-Leonardi 2024).

If we ignore the cognitive assumptions of relevance theory, and the foregrounding of materiality in the constraint-based view, we can see that the two frameworks in fact share a core set of claims: the meanings are not given, but need to be constructed in individual communication situation (in a way that in human interaction we would not hesitate to describe as "participatory", in the sense of De Jaegher and Di Paolo 2007; Di Paolo, Cuffari, and De Jaegher

2018). But that construction is enabled and constrained by a broader set of structures—both syntactic, i.e., the structures of individual signals, and conventional, i.e., the interrelations of those structures (in a way that Deacon's account captures)—which are self-referential with regard to the signals. And ultimately, the meanings of those signals are what they do to their recipients.

Formalizing constraints

The signal/boundary framework proposed by John Holland (2012, as discussed in subsection 4.1.4) offers a modelling formalism which can integrate the self-referential causal relationships. This is particularly because Holland's proposal relies on classifier systems as the computational model of agents. Classifier systems (emerging from Holland's earlier, pioneering work on genetic algorithms, see Holland 2010) are a Turing complete model of computations based on if-then rules, typically designed (through evolutionary algorithms trained under supervison or through reinforcement learning) to predict the future behaviour of the system modelled. The rules, more precisely, are intended to map state values to predictions, with some ability to generalize (through "wild cards", i.e., explicitly ignored properties of the state).

While classifier systems might be a highly simplistic model of causal relationships in general, the fact that they explicitly include counterfactual relationships (i.e., can be used to model interventions, in the sense of Woodward 2004, see also section 3.5.1) and that they can account for abstractions, gives them greater expressive power compared to, for example, the classification/infomorphism setup proposed in Barwise and Seligman's account (see subsection 4.1.3).

Further, the direct inclusion of probabilistic considerations and networked and hierarchical structure of agent's interactions in Holland's signal/boundary account marks parallels between the current theoretical account, and Holland's modelling framework. The "tags", i.e., the specifications of conditions for rules' applicability, can serve as an account of the syntactic structure. Together with the urn model of probability, this offers an interesting formal description for the notion of constraint. Urn models capture probability in terms of randomly drawing balls from an urn. In Pólya urn models, in particular, each observation makes the next such observation more likely, as each time a ball of a particular colour is drawn, it is returned to the urn with an additional ball of that colour. Connecting them to classifier systems, Holland considers urn models with entry and exit conditions, which limit the "colours" that can be drawn from the urn (Holland 2012, pp. 135-139). This model allows him to describe boundaries—a special case of constraints.[27] While the application to physical boundaries

---

[27] Interestingly, urn models can be also used to describe non-stationary processes of diffusion. In particular, the urn models have been extended into generalised Polya processes (see Barraza et al. 2025): this model has a very special setup, in that it is non-stationary (non-ergodic), strongly correlated and non-homogenous, as well as Markovian and self-similar. While generalized Polya processes have not been used (to my knowledge) to describe biological systems, they have been used to model various cases of complex adaptive systems outside of the current domain, such as

offers an intuitive mapping between model and target properties (e.g., a cell is an urn and signal molecules are balls), the generalized case abstracts away from the mechanism through which constraints act. This is an idealization, to some degree—as any formal description would be. Nevertheless, the signal/boundary framework captures many of the postulated features of biological communication. As such, it is an intriguing possibility for further developments of the current account—and an avenue for translating the insights emerging from the current project into more precise mathematical models of biological processes.

Naturalization of content reconsidered

Going back to the debate on naturalizing semantics that has opened this chapter, while the results of semantic search do not on their own allow us to compare the adequacy of the various accounts of content, the close reading of examples does bring some insights. The emerging picture of biological semantics, across the organism- and cell-level, reinforces the importance of the receiver in establishing meaning (while the inputs and sender's intention are necessary, the emphasis is on what the signal does to the receiver, *contra* Neander). While "information" is mentioned by some of the results, in the majority of the cases this is not a technical, information-theoretic notion of information. Indeed, I found only two examples where mutual information (as Piccinini's account would require) is invoked explicitly to relate signals and their meanings:

> "The cost C measures the mutual information between the symbols and the meanings, which is the entropy reduction due to the symbol-meaning correlation in the code matrix." (16956183, 104)

> "Specifically, we measured the mutual information between the stimulus and the responses." (252684986, 100)

Out of the remaining five accounts discussed, each one seems to capture at least some of the characteristics emerging in the data. Millikan's biosemantics and Kolchinsky and Wolpert's causal-necessity approach correctly center the receiver and its operations, as influenced by the signal. Isaac's S-vector semantics provides a probabilistic framework to formalize the constraints the signals introduce (without introducing the more specialist formalisms of hierarchy or category theory—though leaving room for them, if needed)—and captures well the in-between role of "information" at the organism level. Finally, Shea's varitel semantics and Miłkowski's correspondence theory account for the requirements of structure—especially Miłkowski's approach to correspondence allows for the self-referentiality that we saw in biological semantics. Importantly, Miłkowski's account is perhaps the only one that explicitly aims to capture the dynamic,

---

epidemic dynamics (Barraza, Pena, and V. Moreno 2020) or software reliability (S. Li, Dohi, and Okamura 2023).

enacted (or participatory) characteristics of meaning (see Miłkowski 2025): most prefer a stable, atemporal view even when considering imperative (or instructive in Floridi's terms), rather than descriptive, meanings (in the way that Millikan 1995, does).

The ecological account of communicative meanings that I'm building in this chapter does not provide an independent alternative to the teleosemantic approach. In fact it requires some additional theoretical apparatus—especially the notion of function, that is a topic of an ongoing philosophical debate, with teleosemantics being perhaps the most successful view in this discussion (a relevant notion of function might also emerge from the organizational account of biological systems, see A. Moreno and Mossio 2015).[28]

Instead, the view put forward here suggests the direction that these accounts should aim to develop in, in order to accurately capture the actual conceptual practices of life and mind sciences: by connecting the relevance of signals for receiver's operation, through their structural properties that act dynamically within a large system of structures. For such signals, grounding becomes relative to the set of structures under consideration and in that sense perspectival. But since at lower levels of biological organization these structures are material, there is a "ground truth" which secures the meanings.

---

[28] I have limited the discussion of biological functions to minimum in this chapter, as the philosophical literature on this topic is vast and including the topic would go well beyond the scope of the thesis. However, as I believe to have shown, the proposed account of semantics is agnostic with regard to the particular understanding of function, hence this decision should not detract from the current proposal.

6

CONCLUSIONS AND OUTLOOKS

The goal of this thesis was to develop an account of BIOLOGICAL COMMUNICATION across life and mind sciences, covering all the different spatiotemporal scales at which these disciplines employ communicative notions. The initial observation, which motivated this project, was that the concept recurs from the lowest levels of biological organization—subcellular molecular interactions—to some of the highest and most complex—human language and other social phenomena, in particular. This observation motivated the hypothesis that COMMUNICATION has the same meaning and function across the different scales: namely, that communicative processes are scale-free—that they have an underlying organization that can be instantiated at different scales, and that at all those scales they serve a similar epistemic function for biological and cognitive research.

To explore this hypothesis, I have conducted a text mining study over a large corpus of approximately 1.1 million scientific articles from various disciplines of biology and psychology, using the methods of digital philosophy of science. The study has confirmed that a general definition-schema, which I have theoretically derived from the philosophical discussions of animal communication applies across the scales of biological organization. Animal communication is where the concept has been most extensively covered, despite a general lack of agreement over which properties are definitional for communicative interactions (information or influence). The schema postulates three central components:

(1) transfer of a *certain mark* K,
(2) activity of *distinguishable* entities, X and Y, and
(3) the presence of a *structure* S enabling such transfer or exchange.

Note that this is not a complete definition—depending on one's theoretical preferences, the notions of "mark" and "structure" can be pinpointed in different ways (allowing for both informational and influence-based definitions to be coherent with this schema). Most importantly, this description indicates the causal organization of communicative processes. This, in fact, is one of the most important conclusions from the corpus study: despite theoretical preference to regard communication as a somewhat disembodied process (largely due to the theoretical assumptions related to the widespread understanding of language in such abstract terms), the actual scientific use of the concept consistently notes its causal character, and with that—the particular physical implementation of communicative instances.

The three components of the definition-schema require further specification to appropriately distinguish communication from other causal phenomena, as is the case in the scientific use of the term. Hence, I have considered a number of boundary cases to highlight how the conceptual practice of life and mind science distinguishes communication, which led to a number of further requirements that the components of BIOLOGICAL COMMUNICATION proper must meet:

1. The mark, K, is a physical cause or a sequence of causes, characterized by material discontinuity at least at the point of reception (i.e., when reaching the entity Y; as characterized in subsection 3.5.1).
2. The distinguishable entities, X and Y, actively participate in the process, which requires them to have some degree of detachment or organizational closure (as characterized in subsection 3.5.1 and subsection 3.5.2):
    a. X, the sender, actively produces the mark;
    b. Y, the receiver, actively produces some response in result.
3. The structure, S, involves:
    a. the anticipative potentiation (understood in terms of sensitivity to response) of the sender, X (as discussed in subsection 3.5.1);
    b. energetic asymmetry between the signal and the resulting response, which specifies a direction (relative to a particular communication instance; see the discussion in subsection 3.3.1 and subsection 3.5.2);

c. when communication is semantic, the structure has a particular systemic, self-referential character, which allows the signals to act as constraints on the recipient processes (as discussed in subsection 5.5.1 and 5.6).

1. An individual instance of communication is embedded within a larger causal process, and as such it can be characterized as performing some function in that process (as discussed throughout chapter 5).

With this notion of communication at hand, I have proceeded to consider its relation to information theory and the possibility of semantics of biological signals across the scales. The information theoretic perspective is relevant not only due to the debates surrounding the definition of communication. Indeed, the formalism itself has significantly impacted the developments of biology and is responsible for the emergence of cognitive science as a separate area of research. However, analysing philosophical literature on information throughout the thesis, I have noted important confusions which surround the use of information-theoretic notions. Ultimately, there are several theoretically distinct notions of information which share a common formal core, but differ significantly in the epistemic role they may play in science. Two main approaches can be distinguished in terms of "information transmission" and "information content".

Disentangling these confusions provides a better grasp of how information figures in BIOLOGICAL COMMUNICATION, responding to the philosophical worries discussed in section 1.2. It is the original Shannonian notion of information—the concept of information transmission— that is central to capturing and making sense of the causal structure outlined previously. This is the pattern of use that emerges from the corpus data. For this reason, I considered some important theoretical limitations of this view and argue ultimately that information-transmission perspective should be interpreted realistically in the context of biological communication (subsection 4.2.3).

The understanding of meaning of biological communication is particularly intriguing, as the notion of semantic information has been a major topic of debate in philosophy of mind (in the context of naturalizing mental content) and philosophy of biology (in the context of understanding genetics). Nevertheless, it is common for scientists to be weary of ascriptions of meaning to the processes at, e.g., molecular level—despite the evidence to the contrary, emerging in this study, that such semantic practices are indeed persistent in biological practice. Exploration of the semantics of communication across the many contexts where the concept is used, through a computational study of the corpus, introduces relevant nuance to the view of communication proposed earlier.

Indeed, meaning is not scale-free: there are significant differences in how the scientists view semantics of signals at the organism- and cell-level. I draw on a large body of research in naturalistic semantics to make sense of those different uses (section 5.4 and 5.5). Ultimately, I arrive at a view which grounds the signals at organism-level in the internal processes of individual organisms—including, perhaps most naturally, their cognitive operations. Meanwhile, the meaning of cell-level signals is established in terms of how these signals act as physical constraints on other processes. Despite quite different theoretical assumptions and toolboxes, these two views are not inherently at odds. Towards the end of the chapter I gesture at recent research in ecological psychology which can ultimately explain how these simpler, cell-level semantics develop into the more complex, organism-level form. And analogously—how these complex, organism-level ascriptions can be implemented at lower biological scales. This is an open area of research and I hope that the characterization I offer here can inform further analyses.

While this view of semantics shows that there are indeed qualitative differences in communicative processes at different scales, it does not indicate that these processes are substantially different, or different in kind. Eventually, the analysis in the dissertation supports the starting hypothesis of scale-free communication. The primary epistemic function of the notion is to identify and describe a particular causal organization of biological interactions across the different scales where the term is employed. Before concluding, I want to touch upon two further issues. First, I want to consider the fine line between the descriptive and normative goals of philosophy of science I have attempted to walk, a topic I have bracketed in Introduction and left for the current chapter. Second, I want to explore if the corpus analysis offers any kind of support with respect to the speculative considerations motivating my interest in the topic of communication: namely, what role communication, as a scale-free causal process can play for the problem of scaling-up in the basal cognition research programme, and more broadly—what are the epistemic merits of the broadly biosemiotic perspective in this context, which would view signs and codes as definitional to all life.

6.1 THE NORMATIVE OUTCOMES OF A DESCRIPTIVE PROJECT

One implication of the "practice turn" in philosophy of science (see section 2.1) is the shift away from the normative focus which dominated philosophy of science previously, and has come to be viewed as "typically formulated *ex cathedra* and […] thus criticized for being disconnected from and peripheral to the empirical reality of scientific practice" (Kaiser 2019, p. 37). This is an important consideration, as many commentators view the normative line of inquiry as a distinguishing feature of philosophy, both if philosophy is seen as continuous (Varga 2021) or strongly detached from natural sciences and their field of inquiry (A. Thomasson 2015; A. L. Thomasson 2017). Indeed, some regard the rejection of that view—and the broader claim that natural sciences and philosophy are on a spectrum of inquiry—as *scientism*, "the disposition to regard the natural sciences as the only true channel of knowledge" (Dummett 2010, p. 35), and a significant risk to philosophy's epistemic standing, as proponents of scientism would see philosophy's "task reduced to that of adding ornamentation to the theories of the scientists" (Dummett 2010, p. 35; for an example of rejection of philosophy see Hawking and Mlodinow 2010).

Such charges are likely exaggerated (see the challenges raised in Varga 2021), but the normative perspective remains central in philosophy of science, including the work focused on scientific practice (see Kaiser 2019): whether its philosophical treatment of the various norms impacting the practice (in particular, social norms), using descriptive results to fuel normative considerations regarding how science should proceed, or a host of background assumptions, e.g., "about what are good examples and about how to assess the success of science" (Kaiser 2019, p. 37).

Marie Kaiser (2019) and Somogy Varga (2021) offer two independent, though partially overlapping categorizations of these normative goals. Kaiser distinguishes metanormativity, the claims about how science *should be*, methodological normativity, which concerns how philosophy should interpret the claims of science, and object normativity, which emerges from philosophy of science study of normative objects—e.g., epistemic and social norms present in science. Varga, in turn, identifies three levels of normativity: the first level concerns the aims of a particular field of science, the second—epistemic and practical norms guiding scientific inquiry, which explain how scientific knowledge is acquired and why is it reliable (these two levels largely correspond to Kaiser's object normativity), and the third—scientific concepts, allowing for a philosophical evaluation and revision (this aligns Varga with the conceptual engineering project, see Cappelen (2018), cf. Rorot and Miłkowski (2024); the third level partially corresponds to Kaiser's metanormativity).

Indeed, as I have already suggested in the methodological discussion in chapter 2, I view the descriptive goal of the current project as an important one, and I do develop the notion of BIOLOGICAL COMMUNICATION with the explicit aim to align it with the scientific practice. In this regard, the thesis is not intended as a project of conceptual revision (in the vein proposed by Varga, and more broadly—by the conceptual engineering project). However, while the normative stance has not been formulated explicitly, I do not think that the current thesis presents a solely descriptive project. Indeed, in a recent discussion of the normative implications of digital philosophy of science, Miłkowski and Nowakowski (2025) make a case that a "digital discourse analysis", as they term the methodology analogous to the one used in this work, can act to directly reveal "what norms researchers espouse [and] how they reason about them" (p. 9).

In various places throughout the thesis, I have attempted to disentangle conceptual confusions which have accrued around the framework of communication in biology and in cognitive science. While not deeply revisionary, as the analysis is driven by the textual evidence from the corpus, the conclusions I arrive at especially in chapter 4 and 5 depart from explicit conceptual considerations within science. In a sense, the normative component of the current project is to offer an interpretative frame within which the conceptual practice of science is founded on secure epistemic foundation, and can offer greater understanding of phenomena.

Understanding is increasingly recognized as a major and independent aim of science (aside from the more traditional triad of explanation, prediction, and control). Traditionally, understanding has been viewed as dependent on "grasping a correct explanation" (Strevens 2013; see also Elgin 2007; Khalifa 2017; Potochnik 2017). However, there is an emerging strong pluralistic alternative, which admits that explanation and prediction may be required for a complete understanding (Dellsén 2016), but that it is a broader notion. For instance, Wilkenfeld (2019) argues for a graded notion of understanding, where greater understanding consists in the ability to generate more useful information from an accurate, more minimal representation.[1] Scientific concepts, inasmuch as they have a representational component, can enhance our understanding (so viewed), which gives a naturalistically focused philosophy an important role in contemporary scientific landscape.

Eric Schliesser (2019; Schliesser 2024; see also Lewens 2014) proposed the term "synthetic philosophy" to capture the "style of philosophy that brings together insights, knowledge, and arguments from the special sciences with the aim to offer a coherent account of complex systems" (Schliesser 2019, p. 1). He discusses the work of Daniel Dennett (in particular, Dennett 2017) and Peter Godfrey-Smith (in particular, Godfrey-Smith 2018) as exemplars of this approach: a "post-analytic" (Schliesser 2024, p. 243) reflection, distinct from philosophy of science in its direct engagement with empirical results of special sciences—in its focus on the world, rather than on science. In a similar vein, Thomas Pradeu and his group have been championing the program of "philosophy in science", which

---

[1] See also the analysis in Rorot and Miłkowski (2024), which connects this point of view to the earlier proposals of the Lviv-Warsaw School (e.g., Dąmbska and Dąmbska 1975; Jadacki 1990) and argues for the importance of the predictive component of understanding so viewed.

groups philosophers who "use philosophical tools to address scientific problems and provide scientifically useful proposals" (Pradeu et al. 2024, p. 376).[2]

While the current dissertation does not fit neatly neither within the synthetic philosophy framework, nor the program of philosophy in science, due to its primary focus *on* science, these two proposals couch the positive and normative component of my proposal: the "unificationist" aspect of my treatment of BIOLOGICAL COMMUNICATION allows for integrating knowledge from various domains, offers an epistemic justification for the sharing of methods across the scales, and is intended to offer a greater understanding of the characteristics of biological processes.

It is in this perspective that a possible counter-argument should be viewed: a central descriptive step that I have taken early on in this project is to take scientists' ascriptions of communication seriously. Uncovering this assumption a critic could argue that this realist reading of the communicative language is unmotivated, and that communication might well be "scale-free", but nothing more (in at least some of the cases discussed in this work) than a metaphor or an instrumentalist posit which serves some particular cognitive function, but does not capture the real structure of a particular biological process. A critic could then go on to argue that the notion of communication in biology should be abandoned, or at least ultimately deflated, as the epistemic benefits it may currently offer can be achieved without this term—and without the associated risks of anthropomorphism, etc.

It is true that the realist reading of communication is an assumption of the current work, and as such the analysis presented here cannot provide conclusive arguments (more generally—it is unlikely that empirical evidence can weigh in on such questions). Nevertheless, I do believe that it has firmer grounds than any anti-realist readings. It is unlikely that a fictionalist or instrumentalist with regard to biological communication should argue that the notion of "communication" is *never* to be understood literally. Indeed, part of the counter-argument suggested above consists in viewing communication as an *anthropomorphizing* notion. Hence, even a fictionalist would likely agree that communication is literally applicable at least to the human case. But then, the critic would be responsible for drawing a line between the justified, literal uses of "communication" and metaphorical mentions of the term.

The empirical analysis that this dissertation is based on, indicates that there is no clear, arbitrary line for such a distinction, unless the critic intends to view communication as inherently semantic (see the distinction between organism- and cell-level meaning discussed in chapter 5). However, such a strong, seman-

---

[2] In fact this approach has been promoted earlier, though outside of the biological context, by the Copernicus Center for Interdisciplinary Studies in Krakow (currently a unit of the Jagiellonian University). Copernicus Center has been founded by Michał Heller, a cosmologist and philosopher, and its members include philosophers Bartosz Brożek and Mateusz Hohol. As the Center website notes, "philosophy in science" has been Center's motto from its informal beginnings in the 1970s (see https://www.copernicuscenter.edu.pl/en/about-us/, accessed November 30, 2025).

tic notion of communication will most certainly run into significant trouble when applied in the context of animal communication research—which, as I have underscored throughout, partially aims for a naturalistic explanation of complex forms of human linguistic communication. The strong notion would, however, presuppose those complex forms, leading to circularity. For a non-semantic notion of communication, any delineation will be arbitrary and difficult to defend—a problem which the realist reading entirely avoids.

Further, despite the descriptive foundations of the realist assumption, I believe that this perspective is necessary to attain the more normative goal just discussed. Indeed, if the analysis of COMMUNICATION is to enhance our understanding of life and mind, it is vital to take this central term as directly applicable to the processes it describes in biological and cognitive research. Any anti-realist reading, in denying the connection of "communication" to real-world processes, limits its scope in integrating scientific research, precluding the "synthetic" position I have outlined above.

Finally, though this is in no way an argument here, the speculative point I will turn to now is only possible if we interpret ascriptions of BIOLOGICAL COMMUNICATION realistically.

6.2 SCALE-FREE COMMUNICATION AND THE EMERGENCE OF LIFE AND COGNITION

As I have noted in the introduction, aside from the main goal of this thesis, which is to make sense of BIOLOGICAL COMMUNICATION as a conceptual framework in biological and cognitive sciences, I am also motivated by a more speculative hypothesis. The work is embedded in the context of emerging suggestions that communication and information might offer a unifying framework for all biology (Fields and M. Levin 2020). Indeed, some take a more metaphysical stance which views communication as (one of) the definitional traits of life (Barbieri 2008; Witzany 2019, 2020), or—which is more than a difference of emphasis—would argue for viewing "semiosis and interpretative processes [as] essential components in the dynamics of natural systems" and use this conceptual toolbox to "explanatorily reintegrate living creatures [...] into the natural world," from which reductionist or vitalist views would seek to exclude them (Hoffmeyer 2008, p. xiv).

While the Scale-Free Communication account could be read in the context of the metaphysical point Günther Witzany (2019; 2020) formulated most strongly, for now I want to undertake a slightly less demanding task and consider its explanatory implications in the vein suggested by Jesper Hoffmeyer (2008), in particular by looking at potential explanatory benefits Scale-Free Communication may offer for the emerging basal cognition (Lyon 2020; Lyon et al. 2021) research programme.

Basal cognition views cognition as a genuinely biological phenomenon, and argues that this fact should figure more centrally in the methodological approaches in cognitive sciences, including the selection of systems to study. Its proponents argue that

> the cognitive sciences should at last join the rest of the life sciences in the way they approach their quarry. This involves three essential steps. First, start with the *smallest and simplest organisms* that display the phenomenon of interest (the function, the mechanism). Second, in those organisms *identify principles* from observed and measured patterns of genetic, epigenetic and behavioural interactions. Third, scale up to *more complex organisms* and observe where the similarities and differences actually lie, not simply where we think they must lie. [Lyon et al. (2021), p. 1; emphasis original]

As such, basal cognition introduces two central innovations into the study of cognition:[3] first, it argues for an extension of cognitive sciences' target to organisms without a central nervous system, or without a nervous system at

---

[3] Both innovations have important historical precursors—proponents of basal cognition often underscore the continuity between their approach and that of Charles Darwin and other important figures (see e.g., Lyon (2020), p. 415-416; Lyon et al. (2021), p. 3) as a rhetoric device to strengthen their proposal.

all; second, it argues for an overhaul of the field's methodology—for cognitive science to go "fully evolutionary". The methodological arguments were recently reviewed by Alejandro Fábregas-Tejeda and Matthew Sims (2025; see also Sims 2024), and while they will come up in the following discussion, I will focus primarily on the first point.[4]

Basal cognition (particularly as framed by Pamela Lyon, see Lyon 2006; Lyon 2015; Lyon 2020; and Michael Levin, see M. Levin 2019; M. Levin 2022; M. Levin and Dennett 2020; see also Lyon et al. 2021) follows the "biogenic" approach to cognition, as distinguished from the more traditional, "anthropogenic" perspective (Lyon 2006). The latter proceeds by taking certain human capacities as paradigmatic for cognition and proceeds "down" for a broader view, looking for their instantiations or analogues in other species. Instead, the biogenic perspective views cognition as an inherently biological characteristic, and attempts to offer a "phyletically-neutral" definition:

> Cognition is comprised of sensory and other information-processing mechanisms an organism has for becoming familiar with, valuing, and interacting productively with features of its environment in order to meet existential needs, the most basic of which are survival/persistence, growth/thriving, and reproduction. (Lyon 2020, p. 416)

In a similar vein, Michael Levin (2022, p. 15) follows the characterization of "intelligence" proposed by William James, who claimed that for "intelligent agents, altering the conditions changes the activity displayed, but not the end reached; for here the idea of the yet unrealized end co-operates with the conditions to determine what the activities shall be" (James 1981, p. 21). James focused in that statement on the role of goals, final causes, in shaping actions. Levin operationalizes his proposal into a view of cognitive behaviour which underscores the ability of the systems to achieve "fixed ends with varying means." Indeed, such description fits not only a rat navigating around a novel obstacle in a maze, but also the cells of a tadpole during development. For the latter, consider that the tadpole is able to shift facial features around and achieve a normal target morphology even if a surgical procedure scrambles the starting conditions (as Levin's group has shown, e.g., Vandenberg, D. S. Adams, and M. Levin 2012).

Hence, motivated by a strong, Darwinian gradualism with regard to cognition (cf. Cisek 2019, 2022; Cisek and Hayden 2022), the basal programme attempts to identify instances of genuine cognition—at least according to the definitions above—at the smallest and simplest possible level, namely that of unicellular organisms. Bacteria (Gershman et al. 2021; Wan and Jékely 2021; Lyon 2015) and slime moulds (Boussard et al. 2021; Sims 2025) feature most prominently. More complex organisms are also of interest to the proponents of basal cognition, but as I have mentioned, they centre organisms with simple nervous systems—if

---

[4] A more thorough discussion of the basal cognition programme can be found in Harrison, Rorot, and Laukaityte (2022), Falandays, Kaaronen, et al. (2023), and Rorot (under review).

they have one at all (e.g., plants, Baluška and Mancuso 2021; Calvo Garzón and Keijzer 2011; Segundo-Ortin and Calvo 2022; or the *Hydra*, which has a simple nerve net, see Hanson 2021).

An important corollary of that view is a wide-ranging deflationism with regard to standard cognitive or mental categories. Psychological notions, such as BELIEF, DESIRE, or INTENTION, will be viewed as configurations of internal states that can be ascribed to a broad range of biological systems. For instance, one line of research focuses on valence (Damasio 2000; Lyon and Kuchling 2021; Man and Damasio 2019; see also Hodges and Baron 1992; Hodges and Rączaszek-Leonardi 2021), and attempts to tie it to homeostatic and allostatic behaviours, suggesting that the bodily underpinnings of affect (and DESIRE) may have been already present in unicellular life. Similarly, BELIEF, understood as information about some internal or external states, and INTENTION, goal-directed activity based on "beliefs" so-understood, is in this context seen as a feature of biological systems across scales.

Michael Levin and Daniel Dennett (2020) postulate the view of "cognition all the way down", which refers not only to viewing unicellular life as exhibiting purposeful activity, but also individual cells making up multicellular organisms as "agents with agendas". One of the implications of such a view is that we can attempt to model the activity of the larger, multicellular systems as coordinated collective behaviour of their simpler components (see also Falandays, Kaaronen, et al. 2023). In this context, the conceptual reliance of relevance theory on the notion of INTENTION (a topic discussed in chapter 5) can be reformulated into such simpler categories. For instance, we could attempt to describe bioelectric and chemical signalling during development through the "informative intentions" (in the sense of Heintz and Scott-Phillips 2023) of sender cells to let their neighbours know what is their mutual location with respect to the target morphology, or what direction of differentiation they take. What this view achieves is that it underscores the role of information available to the sender cell in establishing the signal's meaning. This connects closely to some crucial research areas, including the study of bioelectric communication in carcinogenesis and tumour proliferation (e.g., Carvalho 2021).

While the basal cognition programme has received significant criticisms regarding the conceptualizations adopted (F. Adams 2018; see the response in Sims 2021), a more important line of criticism suggests that the evolutionary distance between target systems of basal cognition from human cognition renders any evidence and conclusions of the programme essentially irrelevant to the scientists interested in the human case. I.e., that the "scaling-up" postulate of basal cognition is unattainable, regardless of the merit of studying cognition in such simple organisms.[5]

---

[5] The point was raised by John Krakauer in a discussion after Pamela Lyon's talk at the School of Ideas in Neuroscience 2023 in Warsaw. For him, the problem is not limited to the novel target systems basal cognition introduces, but applies also to some of the more established

These critics are correct in identifying a gap in basal cognition research. Despite the methodological postulates and some initial suggestions, the majority of the connections from the simplest systems the basal programme targets to the more complex which are the traditional focus in neuroscience so far is at best a matter of analogy.[6] A thorough examination of the evolutionary connections of the mechanisms (following the methodology outlined by Cisek 2022), or even a theory of the principles of the cognitive "scaling up", is still lagging behind.

The emerging view of BIOLOGICAL COMMUNICATION can aid with explaining the principles of cognitive "scaling up" (for a related point see M. Levin 2023). I do not wish to claim that cognition is communication or anything of the sort (as this overlooks the crucial energetic underpinnings of cognition and life, realized by active matter systems, see Harrison 2023; Man and Damasio 2019; Seth 2025). Instead, the study of communicative organization of cognitive systems offers a way to narrow down the focus when looking for organizational principles that cognitive systems have come to employ in the course of evolution, to expand the scope of the information and action available to them.[7] The results of this dissertation lead to a view of communication as a materially-grounded causal process. This process can be (at least partially) adequately described in terms of information transfer (without becoming disembodied), and is inherently meaningful—precisely through how its physical makeup impacts, shapes and constrains further activities of the systems involved.

Communication so viewed is powerful enough to enable an integration of smaller-scale cognitive phenomena into a larger cognitive process. There's some initial empirical data to support this claim. For instance, Levin's work on bioelectric signalling is largely driven by a closely related view (for recent review, see M. Levin 2023), highlighting how a signalling structure integrates cells in a developing organism, and how modifying that structure impacts the resulting morphology. Meanwhile, the study of collective intelligence explores what

---

neuroscientific animal models, such as *C. elegans*. I am not aware of any paper that develops this criticism further.

Carrie Figdor, in a series of papers (2022; 2024a; 2024b) and during her presentation at the annual meeting of British Society for Philosophy of Science in 2024, makes a closely related point but frames it as an important open research question for basal cognition and related projects, rather than as a damning critique, as Krakauer does. Figdor underscores the importance of distinguishing between genuine homologies and instances of analogy or convergent evolution, necessary for attaining the epistemic goals of what she terms the "phylogenetic" view of cognition. A different line of reasoning in response to this challenge is considered in Rorot (under review).

[6] Consider the reasoning in Hanson (2021), where she focuses on the presence of spontaneuos, low-frequency electric oscillations in various organisms—from bacteria, through *Hydra* and zebrafish, to the default mode network in mammals, including humans: Hanson considers functional similarities, but brackets differences in mechanism.

[7] M. Levin (2019) describes this in terms of a "computational boundary", "the spatio-temporal boundary of events that it can measure, model, and try to affect." (p. 1) and uses the related notion of a "cognitive light cone" to capture what events are cognitively and behaviourally accessible to the system in time or space, its "area of concern" (p. 9). The notion is intended to highlight the possibility of the composition or coupling of individuals into a compound self, which has an emergent, larger light cone than the individuals making up the collective.

patterns of organism-level interaction (including communication) lead to the emergence of cognitive characteristics at supra-individual level (for recent reviews, which also extend the collective perspective to other biological scales, see Falandays, Kaaronen, et al. 2023; McMillen and M. Levin 2024).

In a sense, this is an obvious proposal—especially if we take the informational view on communication, it is clear that the integration of individual activities of components within a cognitive system requires some form of information-sharing between them (even if we take a strongly modular view, there must be a mechanism which orchestrates the activity of the modules and requires some information transfer). But note that the view of communication I propose here is stronger: it requires that such an information sharing be active and characterized by anticipative potentiation to count as communication.

Hence, it requires particular functional capacities on the part of the components making up the communicative system, and a particular causal structure of their interaction. Considering the material embeddedness I have underscored throughout, communication involves a physical integration of the system, which goes beyond what the standard informational view postulates, and cannot be easily described in "merely causal" categories (given the requirement of material discontinuity). In a sense, we can view it as a pathway to identifying "systems" (in the broad sense of General Systems Theory, as problematized by Rosen 1986; see also Bertalanffy 1968).

In that context, this view is narrower than the view of information flow (in the sense of Barwise and Seligman 1997, as I have already outlined in chapter 4) which was also intended to capture an important characteristic of systems—in particular, distributed systems. The difference lies in the way the proposed account of communication replaces Barwise and Seligman's definitional inclusion of "regularities" with stronger claims about causal organization of the system. While both "regularities" and causal connections can be described in terms of classifications and infomorphisms or infocorrespondences, the causal requirements privilege a particular internal perspective on the system, centring relations relevant *for the system*. Even if the system can be characterized through the lens of multiple communication channels, some of these perspectives might be "more real", i.e., more robustly capture genuine causal interactions, than others.[8] This does not require any form of vitalism, it is enough to consider the closure—e.g.,

---

[8] This view is motivated by a process ontology that takes interactions—intra-actions, to use the more technical term offered by Karen Barad (2007)—as an ontological primitive. In that view—which has recently seen some uptake in philosophy of mind and cognitive science (see McGann 2024; see also Dingemanse et al. 2023)—"phenomena" are considered "ontologically primitive relations—relations without preexisting relata" (Barad 2007, p. 333). "Intra-actions" are the "inner goings-on" within the processes that reconstitute the notion of causality, and agency. They enable a causal and agential structure within phenomena: "It is not that a preexisting entity receives a mark [cf. Reichenbach (1958); Salmon (1978); Seibt (2009); —WR] from a separately determinate entity but rather that the marking or specific materializing 'effect' identifies the agencies of observation as agentially separable from its 'cause' (the 'object') within the phenomenon. [...] Cause and effect emerge through intra-actions. Agential intra-actions are causal enactments." (Barad 2007, p. 176)

through recursive self-maintenance—of the syntactic structure of biological communication (required for semantics, see subsection 5.5.1), which enables this internal perspective. Slightly more broadly, though in the same vein, Miłkowski in his correspondence account of semantics (see subsection 5.2.2) ties the relevant infocorrespondences to causal-mechanistic structure of the processes.

The current proposal offers not only a conceptual framing of the "scaling up" of communication, but suggests a possible pathway for developing computational models of increasingly complex communicative systems through the signal/boundary model described in subsection 4.1.3 and section 5.6. Paired with empirical data, this suggests a way for testing the hypotheses presented here. There will likely be a gap between modelling bioelectric signalling in a *B. subtilis* biofilm or in the regenerating *Planaria* tissue, and modelling the interactions in a central nervous system. Nevertheless, the emerging framework allows for further specification of such differences, e.g., in terms of the rule-set complexity of the classifier system involved. Such specification is necessary to connect the basal cognition research to more established neuroscientific programs, offering a well-delineated platform for debate which goes beyond a—likely impossible to resolve—discussion about what cognition is (see also Allen 2017). The communicative frame allows for a pluralism with regard to cognition, but enables also an exchange between the proponents of basal cognition and researchers interested primarily in more traditional targets of neuroscientific study.

It is in the context of the perspectival, integrative pluralism with regard to the notion of "cognition" (see Allen 2017) that I put forward the hypothesis regarding communication as a possible scaling up principle for cognitive systems. The account proposed here provides a piece for the puzzle of integrating the living creatures back into the natural world (Hoffmeyer 2008, p. xiv), by showing how communication—as a natural process—can permeate the biological realm, and how scientists may draw on this notion to describe and explain a variety of complex behaviours of organism—including their intentionality.

I do not have space to argue in greater detail for this view, especially as it is not crucial for the current point—which can be also read through the more standard lens of perspectival realism (Massimi 2022) or through the Wimsattian notion of "robustness" (Wimsatt 2007). I have outlined this metaphysical perspective in greater detail in a talk at the 6th International Conference on Interaction, Language, and Communication in Oslo, Norway in March 2024, and in a working draft titled "Intra-actions of morphogenesis. Towards an "interaction-first" approach to cognition".

BIBLIOGRAPHY

Adams, Fred. 2018. “Cognition Wars.” *Studies in History and Philosophy of Science Part A* 68 (April): 20–30. https://doi.org/10.1016/j.shpsa.2017.11.007.

Afeltowicz, Łukasz, and Witold Wachowski. 2015. “How Far We Can Go Without Looking Under the Skin: The Bounds of Cognitive Science.” *Studies in Logic, Grammar and Rhetoric* 40, no. 1 (March): 91–109. https://doi.org/10.1515/slgr-2015-0005.

Alfano, Mark. 2019. *Nietzsche’s Moral Psychology.* Cambridge: Cambridge University Press.

Allen, Colin. 2017. “On (Not) Defining Cognition.” *Synthese* 194, no. 11 (November): 4233–4249. https://doi.org/10.1007/s11229-017-1454-4.

Allen, Colin, and Marc Hauser. 1992. “Communication and Cognition: Is Information the Connection?” *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1992:81–91.

Allen, Colin, and Jaimie Murdock. 2022. “LDA Topic Modeling: Contexts for the History & Philosophy of Science.” In *Dynamics of Science: Computational Frontiers in History and Philosophy Of.* Edited by Grant Ramsey and Andreas De Block. Pittsburgh: University of Pittsburgh Press.

Anastassiou, Costas A, Rodrigo Perin, Henry Markram, and Christof Koch. 2011. “Ephaptic Coupling of Cortical Neurons.” *Nature Neuroscience* 14, no. 2 (February): 217–223. https://doi.org/10.1038/nn.2727.

Anderson, Michael L. 2015. “Mining the Brain for a New Taxonomy of the Mind.” *Philosophy Compass* 10, no. 1 (January): 68–77. https://doi.org/10.1111/phc3.12155.

Anjum, Rani Lill, and Stephen Mumford. 2018. “Dispositionalism: A Dynamic Theory of Causation.” In *Everything Flows,* 1st ed., edited by Daniel J. Nicholson and John Dupré, 61–75. Oxford: Oxford University Press, May. https://doi.org/10.1093/oso/9780198779636.003.0003.

Ankeny, Rachel A., Hasok Chang, Marcel Boumans, and Mieke Boon. 2011. “Introduction: Philosophy of Science in Practice.” *European Journal for Philosophy of Science* 1, no. 3 (October): 303–307. https://doi.org/10.1007/s13194-011-0036-4.

Ankeny, Rachel A., and Sabina Leonelli. 2016. “Repertoires: A Post-Kuhnian Perspective on Scientific Change and Collaborative Research.” *Studies in History and Philosophy of Science Part A* 60 (December): 18–28. https://doi.org/10.1016/j.shpsa.2016.08.003.

Anscombe, G. E. M. 2000. *Intention.* Cambridge, Massachusetts London, England: Harvard University Press.

Arbib, Michael A., ed. 2003. *The Handbook of Brain Theory and Neural Networks.* Cambridge, Mass: MIT Press.

Archetti, Marco, and Kenneth J. Pienta. 2019. “Cooperation among Cancer Cells: Applying Game Theory to Cancer.” *Nature Reviews Cancer* 19, no. 2 (February): 110–117. https://doi.org/10.1038/s41568-018-0083-7.

Artiga, Marc. 2021. “Signals Are Minimal Causes.” *Synthese* 198, no. 9 (September): 8581–8599. https://doi.org/10.1007/s11229-020-02589-0.

Atkin, Albert. 2023. "Peirce's Theory of Signs." In *The Stanford Encyclopedia of Philosophy,* Spring 2023, edited by Edward N. Zalta and Uri Nodelman. Metaphysics Research Lab, Stanford University.

Auger, Peter. 1998. *Information Sources in Grey Literature.* De Gruyter, December. https://doi.org/10.1515/9783110977233.

Baluška, František, and Michael Levin. 2016. "On Having No Head: Cognition throughout Biological Systems." *Frontiers in Psychology* 7.

Baluška, František, and Stefano Mancuso. 2021. "Individuality, Self and Sociality of Vascular Plants." *Philosophical Transactions of the Royal Society B: Biological Sciences* 376, no. 1821 (March): 20190760. https://doi.org/10.1098/rstb.2019.0760.

Barad, Karen. 2007. *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning.* Durham London: Duke University Press.

Barbieri, Marcello. 2008. "Biosemiotics: A New Understanding of Life." *Naturwissenschaften* 95, no. 7 (July): 577–599. https://doi.org/10.1007/s00114-008-0368-x.

Barker, Peter, Xiang Chen, and Hanne Andersen. 2003. "Kuhn on Concepts and Categorization." In *Thomas Kuhn,* edited by Thomas Nickles, 212–245. Contemporary Philosophy in Focus. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511613975.

Barla, Josef. 2023. "Beyond Reflexivity and Representation: Diffraction as a Methodological Sensitivity in Science Studies." *Distinktion: Journal of Social Theory* 24, no. 3 (September): 444–466. https://doi.org/10.1080/1600910X.2021.1934506.

Barlow, Horace B. 2012. "Possible Principles Underlying the Transformations of Sensory Messages." In *Sensory Communication: Contributions to the Symposium on Principles of Sensory Communication, July 19-August 1, 1959, Endicott House, M.I.T,* edited by Walter A. Rosenblith. Cambridge, Massachusetts: MIT Press.

Barraza, Nestor R, Gabriel Pena, Juliana Gambini, and M Florencia Carusela. 2025. "A Non-Homogeneous, Non-Stationary and Path-Dependent Markov Anomalous Diffusion Model." *Journal of Physics A: Mathematical and Theoretical* 58, no. 9 (March): 095001. https://doi.org/10.1088/1751-8121/adb6df.

Barraza, Néstor Ruben, Gabriel Pena, and Verónica Moreno. 2020. "A Non-Homogeneous Markov Early Epidemic Growth Dynamics Model. Application to the SARS-CoV-2 Pandemic." *Chaos, Solitons & Fractals* 139 (October): 110297. https://doi.org/10.1016/j.chaos.2020.110297.

Barrett, Louise. 2018. "The Evolution of Cognition: A 4E Perspective." In *The Oxford Handbook of 4E Cognition,* 1st ed., edited by Albert Newen, Leon De Bruin, and Shaun Gallagher, 719–734. Oxford University Press, October. https://doi.org/10.1093/oxfordhb/9780198735410.013.38.

Barron, Andrew B., and Jenny Aino Plath. 2017. "The Evolution of Honey Bee Dance Communication: A Mechanistic Perspective." *Journal of Experimental Biology* 220, no. 23 (December): 4339–4346. https://doi.org/10.1242/jeb.142778.

Barsalou, Lawrence W. 1999. "Perceptual Symbol Systems." *The Behavioral and Brain Sciences* 22, no. 4 (August): 577–609, discussion 610–660.

Barwise, Jon, and Jerry Seligman. 1997. *Information Flow.* West Nyack: Cambridge University Press.

Basile, Pierpaolo, Lucia Siciliani, Elio Musacchio, and Giovanni Semeraro. 2025. *Exploring the Word Sense Disambiguation Capabilities of Large Language Models,* arXiv:2503.08662, March. Accessed November 20, 2025. https://doi.org/10.48550/arXiv.2503.08662.

Bateson, Gregory. (1972) 1987. “Form, Substance, and Difference.” In *Steps to an Ecology of Mind.* Northvale, NJ; London: Jason Aronson Inc.

———. 1979. *Mind and Nature: A Necessary Unity.* New York: Dutton.

Bayer, Péter, Robert A. Gatenby, Patricia H. McDonald, Derek R. Duckett, Kateřina Staňková, and Joel S. Brown. 2022. “Coordination Games in Cancer.” Edited by Jun Tanimoto. *PLOS ONE* 17, no. 1 (January): e0261578. https://doi.org/10.1371/journal.pone.0261578.

Bechtel, William. 2016. “Investigating Neural Representations: The Tale of Place Cells.” *Synthese* 193, no. 5 (May): 1287–1321. https://doi.org/10.1007/s11229-014-0480-8.

Bechtel, William, and Andrew Bollhagen. 2021. “Active Biological Mechanisms: Transforming Energy into Motion in Molecular Motors.” *Synthese* 199, nos. 5-6 (December): 12705–12729. https://doi.org/10.1007/s11229-021-03350-x.

———. 2024. “Causal Cycles in Biology.” In *The Routledge Handbook of Causality and Causal Methods,* 1st ed., 255–268. New York: Routledge, December. Accessed February 4, 2025. https://doi.org/10.4324/9781003528937-29.

Beebe, James R., and Finnur Dellsén. 2020. “Scientific Realism in the Wild: An Empirical Study of Seven Sciences and History and Philosophy of Science.” *Philosophy of Science* 87, no. 2 (April): 336–364. https://doi.org/10.1086/707552.

Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. “On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? ▯.” In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency,* 610–623. Virtual Event Canada: ACM, March. https://doi.org/10.1145/3442188.3445922.

Bennett, Jonathan Francis. 1979. *Linguistic Behaviour.* London: Cambridge University Press.

Bergstrom, Carl T., and Michael Lachmann. 2003. “The Red King Effect: When the Slowest Runner Wins the Coevolutionary Race.” *Proceedings of the National Academy of Sciences* 100, no. 2 (January): 593–598. https://doi.org/10.1073/pnas.0134966100.

———. 2004. “Shannon Information and Biological Fitness.” In *Information Theory Workshop,* 50–54. IEEE.

Bergstrom, Carl T., and Martin Rosvall. 2011. “The Transmission Sense of Information.” *Biology & Philosophy* 26, no. 2 (March): 159–176. https://doi.org/10.1007/s10539-009-9180-z.

Berk, Veysel, Jiunn C. N. Fong, Graham T. Dempsey, Omer N. Develioglu, Xiaowei Zhuang, Jan Liphardt, Fitnat H. Yildiz, and Steven Chu. 2012. “Molecular Architecture and Assembly Principles of *Vibrio Cholerae* Biofilms.” *Science* 337, no. 6091 (July): 236–239. https://doi.org/10.1126/science.1222981.

Berry, David M. 2012a. “Introduction: Understanding the Digital Humanities.” In *Understanding Digital Humanities,* edited by David M. Berry, 1–20. London: Palgrave Macmillan UK. https://doi.org/10.1057/9780230371934_1.

———, ed. 2012b. *Understanding the Digital Humanities.* London: Palgrave Macmillan UK. https://doi.org/10.1057/9780230371934_1.

Bertalanffy, Ludwig von. 1968. *General Systems Theory: Foundations, Development, Applications.* New York: George Braziller.

Bertoldi, Nicola, Francis Lareau, Charles H Pence, and Christophe Malaterre. 2024. "A Quantitative Window on the History of Statistics: Topic-Modelling 120 Years of *Biometrika*." *Digital Scholarship in the Humanities* 39, no. 1 (April): 13–29. https://doi.org/10.1093/llc/fqad072.

Bich, Leonardo, and Ramiro Frick. 2018. "Synthetic Modelling of Biological Communication: A Theoretical and Operational Framework for the Investigation of Minimal Life and Cognition." *Complex Systems* 27, no. 3 (October): 267–288. https://doi.org/10.25088/ComplexSystems.27.3.267.

Bickhard, Mark H. 2000. "Autonomy, Function, and Representation." *Communication and Cognition-Artificial Intelligence* 17 (3-4): 111–131.

———. 2007. "Language as an Interaction System." *New Ideas in Psychology* 25, no. 2 (August): 171–187. https://doi.org/10.1016/j.newideapsych.2007.02.006.

———. 2009. "The Interactivist Model." *Synthese* 166, no. 3 (February): 547–591. https://doi.org/10.1007/s11229-008-9375-x.

———. 2020. "Interactivism." In *The Routledge Companion to Philosophy of Psychology,* Second edition, edited by Sarah Robins, John Symons, and Paco Calvo. Abingdon, Oxon ; New York, NY: Routledge, Taylor & Francis Group.

Bickhard, Mark H., and Loren Terveen. 1996. *Foundational Issues in Artificial Intelligence and Cognitive Science: Impasse and Solution.* Amsterdam ; New York: Elsevier.

Bielecka, Krystyna. 2015. "Why Taddeo and Floridi Did Not Solve the Symbol Grounding Problem." *Journal of Experimental & Theoretical Artificial Intelligence* 27, no. 1 (January): 79–93. https://doi.org/10.1080/0952813X.2014.940138.

———. 2024. "Confabulations in Korsakoff's Syndrome: Defending an Error-Based Account." *Erkenntnis* (August). https://doi.org/10.1007/s10670-024-00857-w.

Bielecka, Krystyna, and Marcin Miłkowski. 2020. "Error Detection and Representational Mechanisms." In *What Are Mental Representations?,* 1st ed., edited by Joulia Smortchkova, Krzysztof Dołęga, and Tobias Schlicht, 287–318. New York: Oxford University Press, December. https://doi.org/10.1093/oso/9780190686673.003.0011.

———. 2024. "Representationalism and Rationality: Why Mental Representation Is Real." *Synthese* 203, no. 5 (May): 161. https://doi.org/10.1007/s11229-024-04540-z.

Birhane, Abeba, and Marek McGann. 2024. "Large Models of What? Mistaking Engineering Achievements for Human Linguistic Agency." *Language Sciences* 106 (November): 101672. https://doi.org/10.1016/j.langsci.2024.101672.

Bishop, Mark, and Slawomir J. Nasuto. 1999. *Communicating Neurons: An Alternative Connectionism.*

Björk, Bo-Christer, and David Solomon. 2012. "Open Access versus Subscription Journals: A Comparison of Scientific Impact." *BMC Medicine* 10, no. 1 (December): 73. https://doi.org/10.1186/1741-7015-10-73.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan. 2003. "Latent Dirichlet Allocation." *Journal of Machine Learning Research* 3 (March): 993–1022.

Block, Ned. 1995. "On a Confusion about a Function of Consciousness." *Behavioral and Brain Sciences* 18, no. 2 (June): 227–247. https://doi.org/10.1017/S0140525X00038188.

Boden, Margaret A. 2006. *Mind as Machine: A History of Cognitive Science.* Vol. 1 and 2. Oxford: Clarendon Press.

Bohl, Katrin, Sabine Hummert, Sarah Werner, David Basanta, Andreas Deutsch, Stefan Schuster, Günter Theißen, and Anja Schroeter. 2014. "Evolutionary Game Theory: Molecules as Players." *Mol. BioSyst.* 10 (12): 3066–3074. https://doi.org/10.1039/C3MB70601J.

Bokulich, Alisa. 2014. "How the Tiger Bush Got Its Stripes: 'How Possibly' vs. 'How Actually' Model Explanations:" *Monist* 97 (3): 321–338. https://doi.org/10.5840/monist201497321.

Borg, Emma. 2009. "Intention-Based Semantics." In *The Oxford Handbook of Philosophy of Language,* 1st ed., edited by Ernie Lepore and Barry C. Smith, 250–266. Oxford University Press, September. https://doi.org/10.1093/oxfordhb/9780199552238.003.0012.

———. 2025. "The Problems of Polysemy." In *The Oxford Handbook of Contemporary Philosophy of Language,* 1st ed., edited by Ernie Lepore and Una Stojnić, 287–311. Oxford University Press, January. https://doi.org/10.1093/oxfordhb/9780192856852.013.31.

Borghi, Anna M., Laura Barca, Ferdinand Binkofski, Cristiano Castelfranchi, Giovanni Pezzulo, and Luca Tummolini. 2019. "Words as Social Tools: Language, Sociality and Inner Grounding in Abstract Concepts." *Physics of Life Reviews* 29 (July): 120–153. https://doi.org/10.1016/j.plrev.2018.12.001.

Borghi, Anna M., Ferdinand Binkofski, Cristiano Castelfranchi, Felice Cimatti, Claudia Scorolli, and Luca Tummolini. 2017. "The Challenge of Abstract Concepts." *Psychological Bulletin* 143, no. 3 (March): 263–292. https://doi.org/10.1037/bul0000089.

Borghi, Anna M., Albertyna Osińska, Andreas Roepstorff, and Joanna Raczaszek-Leonardi. 2023. "Concepts in Interaction: Social Engagement and Inner Experiences." *Philosophical Transactions of the Royal Society B: Biological Sciences* 378, no. 1870 (February): 20210351. https://doi.org/10.1098/rstb.2021.0351.

Boussard, Aurèle, Adrian Fessel, Christina Oettmeier, Léa Briard, Hans-Günther Döbereiner, and Audrey Dussutour. 2021. "Adaptive Behaviour and Learning in Slime Moulds: The Role of Oscillations." *Philosophical Transactions of the Royal Society B: Biological Sciences* 376, no. 1820 (March): 20190757. https://doi.org/10.1098/rstb.2019.0757.

Boyd, Richard. 1993. "Metaphor and Theory Change: What Is "Metaphor" a Metaphor For?" In *Metaphor and Thought,* 2nd ed, edited by Andrew Ortony, 481–532. Cambridge (Mass.): Cambridge university press.

Bradbury, Jack W., and Sandra Lee Vehrencamp. 1998. *Principles of Animal Communication.* Sunderland: Sinauer associates.

Braddon-Mitchell, David, and Robert Nola, eds. 2009. *Conceptual Analysis and Philosophical Naturalism.* Cambridge, Mass: MIT Press.

Brancazio, Nick. 2023. "Interactive Agential Dynamics." *Synthese* 201, no. 6 (June): 221. https://doi.org/10.1007/s11229-023-04192-5.

Brandon, Robert N. 1994. "Theory and Experiment in Evolutionary Biology." *Synthese* 99 (1): 59–73.

———. 1997. "Does Biology Have Laws? The Experimental Evidence." *Philosophy of Science* 64 (S4): S444–S457. https://doi.org/10.1086/392621.

Brenner, Eric D., Rainer Stahlberg, Stefano Mancuso, Jorge Vivanco, František Baluška, and Elizabeth Van Volkenburgh. 2006. "Plant Neurobiology: An Integrated View of Plant Signaling." *Trends in Plant Science* 11, no. 8 (August): 413–419. https://doi.org/10.1016/j.tplants.2006.06.009.

Brentano, Franz. (1874) 2015. *Psychology from an Empirical Standpoint.* Abingdon, Oxon: Routledge.

Brette, Romain. 2019. "Is Coding a Relevant Metaphor for the Brain?" *Behavioral and Brain Sciences* 42:e215. https://doi.org/10.1017/S0140525X19000049.

Brillouin, Léon. 1962. *Science and Information Theory.* New York: Academic Press Inc.

Broadbent, Donald E. 1952. "Listening to One of Two Synchronous Messages." *Journal of Experimental Psychology* 44 (1): 51–55. https://doi.org/10.1037/h0056491.

Brooks, Rodney A. 1991. "Intelligence without Representation." *Artificial Intelligence* 47, no. 1 (January): 139–159. https://doi.org/10.1016/0004-3702(91)90053-M.

Brown, Harold I. 1996. "The Methodological Roles of Theory in Science." In *The Scientific Nature of Geomorphology: Proceedings of the 27th Binghamton Symposium in Geomorphology, Held 27 - 29 September 1996,* edited by Bruce L. Rhoads and Colin E. Thorn, 3–20. Chichester: Wiley.

Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, et al. 2020. *Language Models Are Few-Shot Learners,* arXiv:2005.14165, July. Accessed January 27, 2023.

Bruner, Jerome S., Jacquelline J. Goodnow, and George A. Austin. 1956. *A Study of Thinking.* New York ; London ; Sydney: John Wiley & Sons, Inc.

Budi, Indra, and Yaniasih Yaniasih. 2023. "Understanding the Meanings of Citations Using Sentiment, Role, and Citation Function Classifications." *Scientometrics* 128, no. 1 (January): 735–759. https://doi.org/10.1007/s11192-022-04567-4.

Bueno, Otavio. 2016. "Epistemology and Philosophy of Science." In *The Oxford Handbook of Philosophy of Science,* edited by Paul Humphreys. Oxford Handbooks Online. Oxford University Press, January. https://doi.org/10.1093/oxfordhb/9780199368815.013.13.

Bühler, Karl. (1934) 2011. *Theory of Language: The Representational Function of Language.* Edited by Achim Eschbach. Translated by Donald Fraser Goodwin. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/z.164.

Burian, Richard M. 2001. "The Dilemma of Case Studies Resolved: The Virtues of Using Case Studies in the History and Philosophy of Science." *Perspectives on Science* 9, no. 4 (December): 383–404. https://doi.org/10.1162/106361401760375794.

Butlin, Patrick. 2020. "Representation and the Active Consumer." *Synthese* 197, no. 10 (October): 4533–4550. https://doi.org/10.1007/s11229-018-01941-9.

Buzsáki, György. 2015. "Hippocampal Sharp Wave-ripple: A Cognitive Biomarker for Episodic Memory and Planning." *Hippocampus* 25, no. 10 (October): 1073–1188. https://doi.org/10.1002/hipo.22488.

———. 2019. *The Brain from inside Out.* New York, NY: Oxford University Press.

Calcott, Brett. 2014. "The Creation and Reuse of Information in Gene Regulatory Networks." *Philosophy of Science* 81, no. 5 (December): 879–890. https://doi.org/10.1086/677687.

Call, Josep. 2011. "How Artificial Communication Affects the Communication and Cognition of the Great Apes." *Mind & Language* 26, no. 1 (February): 1–20. https://doi.org/10.1111/j.1468-0017.2010.01408.x.

Callebaut, Werner, ed. 1993. *Taking the Naturalistic Turn or How Real Philosophy of Science Is Done: Conversations with William Bechtel.* Chicago: University of Chicago Press.

———. 2013. “Naturalizing Theorizing: Beyond a Theory of Biological Theories.” *Biological Theory* 7, no. 4 (June): 413–429. https://doi.org/10.1007/s13752-013-0122-2.

Calvo Garzón, Paco, and Fred Keijzer. 2011. “Plants: Adaptive Behavior, Root-Brains, and Minimal Cognition.” *Adaptive Behavior* 19, no. 3 (June): 155–171. https://doi.org/10.1177/1059712311409446.

Campello, Ricardo J. G. B., Davoud Moulavi, and Joerg Sander. 2013. “Density-Based Clustering Based on Hierarchical Density Estimates.” In *Advances in Knowledge Discovery and Data Mining,* edited by David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell, Moni Naor, et al., 7819:160–172. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-37456-2_14.

Campello, Ricardo J. G. B., Davoud Moulavi, Arthur Zimek, and Jörg Sander. 2015. “Hierarchical Density Estimates for Data Clustering, Visualization, and Outlier Detection.” *ACM Transactions on Knowledge Discovery from Data* 10, no. 1 (July): 1–51. https://doi.org/10.1145/2733381.

Cao, Yu, Yaping Pan, Hua Huang, Xiangshu Jin, Elena J. Levin, Brian Kloss, and Ming Zhou. 2013. “Gating of the TrkH Ion Channel by Its Associated RCK Protein TrkA.” *Nature* 496, no. 7445 (April): 317–322. https://doi.org/10.1038/nature12056.

Cappelen, Herman. 2018. *Fixing Language: An Essay on Conceptual Engineering.* Oxford, United Kingdom ; New York, NY: Oxford University Press.

Cárdenas, FJ Navarro, and Francisco Javier. 2022. “Metaphysics of Biological Individuality: Arguments for a Pluralist Approach.” *Revista de Humanidades de Valparaíso,* no. 20, 271–290.

Carnap, Rudolf, and Yehoshua Bar-Hillel. 1952. *An Outline of a Theory of Semantic Information.* Technical Report 247. Cambridge, Massachusetts: Research Laboratory of Electronics, Massachusetts Institute of Technology, October.

Cartwright, Julyan H. E., Jitka Čejková, Elena Fimmel, Simone Giannerini, Diego Luis Gonzalez, Greta Goracci, Clara Grácio, et al. 2024. “Information, Coding, and Biological Function: The Dynamics of Life.” *Artificial Life* 30, no. 1 (February): 16–27. https://doi.org/10.1162/artl_a_00432.

Cartwright, Nancy. 1983. *How the Laws of Physics Lie.* Oxford University Press, June. https://doi.org/10.1093/0198247044.001.0001.

———. 2019. *Nature, the Artful Modeler: Lectures on Laws, Science, How Nature Arranges the World and How We Can Arrange It Better.* 23. Chicago: Open Court.

Carvalho, Joao. 2021. “A Bioelectric Model of Carcinogenesis, Including Propagation of Cell Membrane Depolarization and Reversal Therapies.” *Scientific Reports* 11, no. 1 (June): 13607. https://doi.org/10.1038/s41598-021-92951-0.

Casey, William, Steven E Massey, and Bud Mishra. 2021. “How Signalling Games Explain Mimicry at Many Levels: From Viral Epidemiology to Human Sociology.” *Journal of The Royal Society Interface* 18, no. 175 (February): rsif.2020.0689, 20200689. https://doi.org/10.1098/rsif.2020.0689.

Chadarevian, Soraya De, and Harmke Kamminga. 2003. *Molecularizing Biology and Medicine.* Taylor & Francis, September. https://doi.org/10.4324/9780203304235.

Chambers, Christopher B., Ying Peng, Hoang Nguyen, Nicholas Gaiano, Gord Fishell, and Jeffrey S. Nye. 2001. “Spatiotemporal Selectivity of Response to Notch1 Signals in Mammalian Forebrain Precursors.” *Development* 128, no. 5 (March): 689–702. https://doi.org/10.1242/dev.128.5.689.

Chang, Hasok. 2007. *Inventing Temperature: Measurement and Scientific Progress.* Oxford: Oxford University Press.

———. 2017. “Operational Coherence as the Source of Truth.” *Proceedings of the Aristotelian Society* 117, no. 2 (July): 103–122. https://doi.org/10.1093/arisoc/aox004.

Chang, Jonathan, Sean Gerrish, Chong Wang, Jordan Boyd-graber, and David Blei. 2009. “Reading Tea Leaves: How Humans Interpret Topic Models.” In *Advances in Neural Information Processing Systems,* edited by Y. Bengio, D. Schuurmans, J. Lafferty, C. Williams, and A. Culotta, vol. 22. Curran Associates, Inc.

Chao, Hsiang-Ke. 2023. “Three Kinds of the Lotka–Volterra Model Transfer from Biology to Economics.” *Synthese* 202, no. 4 (October): 124. https://doi.org/10.1007/s11229-023-04341-w.

Chemero, Anthony. 2000. “Anti-Representationalism and the Dynamical Stance.” *Philosophy of Science* 67, no. 4 (December): 625–647. https://doi.org/10.1086/392858.

———. 2009. *Radical Embodied Cognitive Science.* Cambridge, Mass: MIT Press.

Cheon, Hyundeuk, and Edouard Machery. 2016. “Scientific Concepts.” In *The Oxford Handbook of Philosophy of Science,* edited by Paul Humphreys. Oxford Handbooks Online. Oxford University Press, May. https://doi.org/10.1093/oxfordhb/9780199368815.013.9.

Chirimuuta, Mazviita. 2018. “Marr, Mayr, and MR: What Functionalism Should Now Be About.” *Philosophical Psychology* 31, no. 3 (April): 403–418. https://doi.org/10.1080/09515089.2017.1381679.

———. 2024. *The Brain Abstracted: Simplification in the History and Philosophy of Neuroscience.* Cambridge, Massachusetts London, England: The MIT Press.

Chis-Ciure, Robert, and Michael Levin. 2025. “Cognition All the Way down 2.0: Neuroscience beyond Neurons in the Diverse Intelligence Era.” *Synthese* 206, no. 5 (November): 257. https://doi.org/10.1007/s11229-025-05319-6.

Chomsky, Noam. 1959. “On Certain Formal Properties of Grammars.” *Information and Control* 2, no. 2 (June): 137–167. https://doi.org/10.1016/s0019-9958(59)90362-6.

Cisek, Paul. 2019. “Resynthesizing Behavior through Phylogenetic Refinement.” *Attention, Perception, & Psychophysics* 81, no. 7 (October): 2265–2287. https://doi.org/10.3758/s13414-019-01760-1.

———. 2022. “Evolution of Behavioural Control from Chordates to Primates.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 377, no. 1844 (February): 20200522. https://doi.org/10.1098/rstb.2020.0522.

Cisek, Paul, and Benjamin Y. Hayden. 2022. “Neuroscience Needs Evolution.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 377, no. 1844 (February): 20200518. https://doi.org/10.1098/rstb.2020.0518.

Clark, Andy. 1996. *Being There: Putting Brain, Body, and World Together Again.* Cambridge, Mass.: MIT Press.

———. 2008. *Supersizing the Mind: Embodiment, Action, and Cognitive Extension.* Oxford University PressNew York, December. https://doi.org/10.1093/acprof:oso/9780195333213.001.0001.

———. 2013. "Whatever next? Predictive Brains, Situated Agents, and the Future of Cognitive Science." *Behavioral and Brain Sciences* 36, no. 03 (June): 181–204. https://doi.org/10.1017/S0140525X12000477.

———. 2016a. "Predicting Peace." In *Open MIND,* edited by Thomas K. Metzinger and Jennifer M. Windt. The MIT Press, May. https://doi.org/10.7551/mitpress/10603.003.0024.

———. 2016b. *Surfing Uncertainty: Prediction, Action, and the Embodied Mind.* Oxford ; New York: Oxford University Press.

Clark, Andy, and David Chalmers. 1998. "The Extended Mind." *Analysis* 58, no. 1 (January): 7–19. https://doi.org/10.1093/analys/58.1.7.

Clark, Andy, and Josefa Toribio. 1994. "Doing without Representing?" *Synthese* 101, no. 3 (December): 401–431. https://doi.org/10.1007/BF01063896.

Clarke, Ellen. 2010. "The Problem of Biological Individuality." *Biological Theory* 5, no. 4 (December): 312–325. https://doi.org/10.1162/BIOT_a_00068.

Cobb, Matthew. 2020. *The Idea of the Brain: A History.* London: Profile.

———. 2021. "A Brief History of Wires in the Brain." *Frontiers in Ecology and Evolution* 9 (December): 760269. https://doi.org/10.3389/fevo.2021.760269.

Colaço, David. 2018. "Rethinking the Role of Theory in Exploratory Experimentation." *Biology & Philosophy* 33, nos. 5-6 (December): 38. https://doi.org/10.1007/s10539-018-9648-9.

———. 2025. "On Consistently Assessing Alleged Mnemonic Systems (or, Why Isn't Immune Memory "Really" Memory?)" *Review of Philosophy and Psychology* (February). https://doi.org/10.1007/s13164-025-00768-x.

Collier, John Donald. 1996. "Information Originates in Symmetry Breaking." *Symmetry: Science and Culture* 7:247–256.

———. 1999. "Causation Is the Transfer of Information." In *Causation and Laws of Nature,* edited by Howard Sankey, 215–245. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-015-9229-1_18.

———. 2008. "Information in Biological Systems." In *Handbook of Philosophy of Science, Volume 8: Philosophy of Information,* edited by Pieter Adriaans and Johan van Benthem, 763–787. Dordrecht; North-Holland: Elsevier. https://doi.org/10.1016/B978-0-444-51726-5.50024-8.

———. 2011. "Information, Causation and Computation." In *Information and Computation,* edited by Gordana Dodig-Crnkovic and Mark Burgin, 89–105. WORLD SCIENTIFIC, June. https://doi.org/10.1142/9789814295482_0004.

Conant, Roger C., and W. Ross Ashby. 1970. "Every Good Regulator of a System Must Be a Model of That System †." *International Journal of Systems Science* 1, no. 2 (October): 89–97. https://doi.org/10.1080/00207727008920220.

Couch, Mark B. n.d. "Causal Role Theories of Functional Explanation." *The Internet Encyclopedia of Philosophy.*

Cover, Thomas M., and Joy A. Thomas. 2006. *Elements of Information Theory.* Hoboken, N.J: Wiley-Interscience.

Cowley, Stephen J. 2011. "Taking a Language Stance." *Ecological Psychology* 23, no. 3 (July): 185–209. https://doi.org/10.1080/10407413.2011.591272.

Cox, D. R. 2001. “Biometrika: The First 100 Years.” *Biometrika* 88 (1): 3–11.

Craver, Carl F. 2007. *Explaining the Brain.* Oxford University Press, June. https://doi.org/10.1093/acprof:oso/9780199299317.001.0001.

———. 2009. “Mechanisms and Natural Kinds.” *Philosophical Psychology* 22, no. 5 (October): 575–594. https://doi.org/10.1080/09515080903238930.

Csillag, Márton, Hamza Giaffar, Eörs Szathmáry, Mauro Santos, and Dániel Czégel. 2025. “From Bayes to Darwin: Evolutionary Search as an Exaptation from Sampling-Based Bayesian Inference.” *Journal of Theoretical Biology* 599 (February): 112032. https://doi.org/10.1016/j.jtbi.2024.112032.

Cummins, Robert. 1975. “Functional Analysis.” *The Journal of Philosophy* 72, no. 20 (November): 741. https://doi.org/10.2307/2024640.

———. 1977. “Programs in the Explanation of Behavior.” *Philosophy of Science* 44, no. 2 (June): 269–287. https://doi.org/10.1086/288742.

———. 1983. *The Nature of Psychological Explanation.* Cambridge, Mass.: MIT Press.

Currie, Adrian. 2015. “Philosophy of Science and the Curse of the Case Study.” In *The Palgrave Handbook of Philosophical Methods,* edited by Chris Daly, 553–572. London: Palgrave Macmillan UK. https://doi.org/10.1057/9781137344557_22.

Czégel, Dániel, Hamza Giaffar, Márton Csillag, Bálint Futó, and Eörs Szathmáry. 2021. “Novelty and Imitation within the Brain: A Darwinian Neurodynamic Approach to Combinatorial Problems.” *Scientific Reports* 11, no. 1 (June): 12513. https://doi.org/10.1038/s41598-021-91489-5.

Czégel, Dániel, István Zachar, and Eörs Szathmáry. 2019. “Multilevel Selection as Bayesian Inference, Major Transitions in Individuality as Structure Learning.” *Royal Society Open Science* 6, no. 8 (August): 190202. https://doi.org/10.1098/rsos.190202.

Dahl, Niklas. Forthcoming. “Meaningless Communication: Towards an Understanding-First Approach.” *Journal of Philosophy.*

Damasio, Antonio R. 2000. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness.* San Diego, CA: Harcourt.

Dąmbska, Izydora, and Izydora Dąmbska. 1975. “W sprawie pojęcia rozumienia.” In *Znaki i myśli: wybór pism z semiotyki, teorii nauki i historii filozofii,* 49–56. Warszawa: Państwowe Wydawnictwo Naukowe.

Davidson, Donald. 1963. “Actions, Reasons, and Causes.” *The Journal of Philosophy* 60 (23): 685–700. https://doi.org/10.2307/2023177.

———. 2005. “Meaning, Truth, and Evidence.” In *Truth, Language, and History,* 1st ed., 47–62. Oxford University PressOxford, February. Accessed April 24, 2025. https://doi.org/10.1093/019823757X.003.0004.

Dawkins, Richard. (1976) 2006. *The Selfish Gene.* Oxford ; New York: Oxford University Press.

De Block, Andreas, and Grant Ramsey. 2022. “Introduction. Tools, Tests, and Data: An Introduction to the New History and Philosophy of Science.” In *The Dynamics of Science: Computational Frontiers in History and Philosophy of Science,* edited by Grant Ramsey and Andreas De Block, 3–15. Pittsburgh, Pa: University of Pittsburgh press.

De Jaegher, Hanne. 2021. “Loving and Knowing: Reflections for an Engaged Epistemology.” *Phenomenology and the Cognitive Sciences* 20, no. 5 (December): 847–870. https://doi.org/10.1007/s11097-019-09634-5.

De Jaegher, Hanne, and Ezequiel A. Di Paolo. 2007. “Participatory Sense-Making: An Enactive Approach to Social Cognition.” *Phenomenology and the Cognitive Sciences* 6, no. 4 (December): 485–507. https://doi.org/10.1007/s11097-007-9076-9.

Deacon, Terrence W. 1998. *The Symbolic Species: The Co-Evolution of Language and the Brain.* New York, NY: Norton.

———. 2007. “Shannon – Boltzmann – Darwin: Redefining Information (Part I).” *Cognitive Semiotics* 1, no. s1 (September): 123–148. https://doi.org/10.1515/cogsem.2007.1.fall2007.123.

———. 2008. “Shannon - Boltzmann — Darwin: Redefining Information (Part II).” *Cognitive Semiotics* 2, no. Supplement (January). https://doi.org/10.1515/cogsem.2008.2.spring2008.169.

———. 2011. *Incomplete Nature: How Mind Emerged from Matter.* New York: W.W. Norton & Co.

———. 2017. “Language Evolution and Neuromechanisms.” In *A Companion to Cognitive Science,* 1st ed., edited by William Bechtel and George Graham, 212–225. Wiley, August. https://doi.org/10.1002/9781405164535.ch13.

———. 2021. “How Molecules Became Signs.” *Biosemiotics* 14, no. 3 (December): 537–559. https://doi.org/10.1007/s12304-021-09453-9.

Deacon, Terrence W., and Joanna Rączaszek-Leonardi. 2019. “Abandoning the Code Metaphor Is Compatible with Semiotic Process.” *Behavioral and Brain Sciences* 42:e224. https://doi.org/10.1017/S0140525X19001419.

Dębowski, Łukasz Jerzy. 2021. *Information Theory Meets Power Laws: Stochastic Processes and Language Models.* Hoboken: Wiley.

Deely, John. 2011. “Thomas A. Sebeok and Semiotics of the 21st Century.” In *Semiotics Continues to Astonish,* edited by Paul Cobley, John Deely, Kalevi Kull, and Susan Petrilli, 123–160. De Gruyter Mouton, July. https://doi.org/10.1515/9783110254389.123.

Deleuze, Gilles, Félix Guattari, Janis Tomlinson, and Graham Burchell III. 2014. *What Is Philosophy?* New York: Columbia University Press.

Dellsén, Finnur. 2016. “Scientific Progress: Knowledge versus Understanding.” *Studies in History and Philosophy of Science Part A* 56 (April): 72–83. https://doi.org/10.1016/j.shpsa.2016.01.003.

Dennett, Daniel C. (1969) 2010. *Content and Consciousness.* Routledge. https://doi.org/10.4324/9780203092958.

———. 1971. “Intentional Systems.” *Journal of Philosophy* 68 (4): 87–106. https://doi.org/10.2307/2025382.

———. 1987. *The Intentional Stance.* Cambridge, Mass.: MIT Press.

———. 1988. “Précis of *The Intentional Stance.*” *Behavioral and Brain Sciences* 11, no. 3 (September): 495–505. https://doi.org/10.1017/S0140525X00058611.

———. 1996. *Darwin’s Dangerous Idea: Evolution and the Meanings of Life.* New York: Simon & Schuster.

Dennett, Daniel C. 2017. *From Bacteria to Bach and Back: The Evolution of Minds.* New York: W.W. Norton & Company.

Descartes, René. (1667) 2018. *L'homme.* Edited by Delphine Antoine-Mahut. 1600. Paris: Flammarion.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,* arXiv:1810.04805, May. Accessed January 27, 2023.

Di Paolo, Ezequiel A., Thomas Buhrmann, and Xabier E. Barandiaran. 2017. *Sensorimotor Life: An Enactive Proposal.* Oxford, United Kingdom: Oxford University Press.

Di Paolo, Ezequiel A., Elena Clare Cuffari, and Hanne De Jaegher. 2018. *Linguistic Bodies: The Continuity between Life and Language.* The MIT Press. https://doi.org/10.7551/mitpress/11244.001.0001.

Dickins, Thomas E. 2023. "The Role of Information in Evolutionary Biology." *Acta Biotheoretica* 71, no. 3 (September): 17. https://doi.org/10.1007/s10441-023-09468-4.

Diener, Marc J. 2010. "Cohen's *d*." In *The Corsini Encyclopedia of Psychology,* 1st ed., edited by Irving B. Weiner and W. Edward Craighead, 1–1. Wiley, January. https://doi.org/10.1002/9780470479216.corpsy0200.

DiFrisco, James. 2023. "Toward a Theory of Homology: Development and the De-Coupling of Morphological and Molecular Evolution." *The British Journal for the Philosophy of Science* 74, no. 3 (September): 771–810. https://doi.org/10.1086/714959.

DiFrisco, James, Alan C. Love, and Günter P. Wagner. 2020. "Character Identity Mechanisms: A Conceptual Model for Comparative-Mechanistic Biology." *Biology & Philosophy* 35, no. 4 (August): 44. https://doi.org/10.1007/s10539-020-09762-2.

———. 2023. "The Hierarchical Basis of Serial Homology and Evolutionary Novelty." *Journal of Morphology* 284, no. 1 (January): e21531. https://doi.org/10.1002/jmor.21531.

Dingemanse, Mark, Andreas Liesenfeld, Marlou Rasenberg, Saul Albert, Felix K. Ameka, Abeba Birhane, Dimitris Bolis, et al. 2023. "Beyond Single-Mindedness: A Figure-Ground Reversal for the Cognitive Sciences." *Cognitive Science* 47, no. 1 (January): e13230. https://doi.org/10.1111/cogs.13230.

Dingemanse, Mark, Seán G. Roberts, Julija Baranova, Joe Blythe, Paul Drew, Simeon Floyd, Rosa S. Gisladottir, et al. 2015. "Universal Principles in the Repair of Communication Problems." Edited by Sonja Kotz. *PLOS ONE* 10, no. 9 (September): e0136100. https://doi.org/10.1371/journal.pone.0136100.

Donaldson-Matasci, Matina C., Carl T. Bergstrom, and Michael Lachmann. 2010. "The Fitness Value of Information." *Oikos* 119, no. 2 (February): 219–230. https://doi.org/10.1111/j.1600-0706.2009.17781.x.

Dretske, Fred I. 1981. *Knowledge & the Flow of Information.* Cambridge, Mass: MIT Press.

———. 1986. "Misrepresentation." In *Belief: Form, Content, and Function,* edited by Radu J. Bogdan, 17–36. Oxford University Press.

———. 1995. *Naturalizing the Mind.* 1994. Cambridge, Mass.: MIT Press.

Dreyfus, Hubert L. 1986. *What Computers Can't Do: The Limits of Artificial Intelligence.* New York: Harper & Row.

———. 1992. *What Computers Still Can't Do: A Critique of Artificial Reason.* Cambridge, Mass: MIT Press.

Dubreuil, Colin, Hugh Notman, Louise Barrett, Peter Henzi, and Mary Susan McDonald Pavelka. 2025. "Sex Differences in the Acoustic Structure of Terrestrial Alarm Calls in Vervet Monkeys (*Chlorocebus Pygerythrus*)." *American Journal of Primatology* 87, no. 1 (January): e23674. https://doi.org/10.1002/ajp.23674.

Dummett, Michael. 1995. *The Logical Basis of Metaphysics.* London: Duckworth.

———. 2010. *The Nature and Future of Philosophy.* New York: Columbia University Press.

Dusenbery, David B. 1996. *Life at Small Scale: The Behavior of Microbes.* 61. New York: Scientific American Library.

Elgin, Catherine. 2007. "Understanding and the Facts." *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 132 (1): 33–42.

Elion, E. 2006. "Methods for Analyzing MAPK Cascades." *Methods* 40, no. 3 (November): 207–208. https://doi.org/10.1016/j.ymeth.2006.10.001.

Elkins, James. 2001. *The Domain of Images.* Ithaca, NY: Cornell Univ. Press.

Emmeche, Claus, and Jesper Hoffmeyer. 1991. "From Language to Nature: The Semiotic Metaphor in Biology." *Semiotica* 84 (1-2): 1–42. https://doi.org/10.1515/semi.1991.84.1-2.1.

European Science Foundation. 2021. *Plan S: Making Full and Immediate Open Access a Reality.* Technical report. Accessed October 4, 2024.

Fábregas-Tejeda, Alejandro, and Matthew Sims. 2025. "On the Prospects of Basal Cognition Research Becoming Fully Evolutionary: Promising Avenues and Cautionary Notes." *History and Philosophy of the Life Sciences* 47, no. 1 (March). https://doi.org/10.1007/s40656-025-00660-y.

Falandays, J. Benjamin, Roope O. Kaaronen, Cody Moser, Wiktor Rorot, Joshua Tan, Vishwanath Varma, Tevin Williams, and Mason Youngblood. 2023. "All Intelligence Is Collective Intelligence." *Journal of Multiscale Neuroscience* 2, no. 1 (April): 169–191. https://doi.org/10.56280/1564736810.

Falandays, J. Benjamin, Jeffrey Yoshimi, William H. Warren, and Michael J. Spivey. 2023. "A Potential Mechanism for Gibsonian Resonance: Behavioral Entrainment Emerges from Local Homeostasis in an Unsupervised Reservoir Network." *Cognitive Neurodynamics* (July). https://doi.org/10.1007/s11571-023-09988-2.

Falk, Jay J., Carl T. Bergstrom, Kevin J.S. Zollman, and Alejandro Rico-Guevara. 2025. "Partial Honesty in a Hummingbird Polymorphism Provides Evidence for a Hybrid Equilibrium." *Animal Behaviour* 222 (April): 123104. https://doi.org/10.1016/j.anbehav.2025.123104.

Feest, Uljana. 2012. "Exploratory Experiments, Concept Formation, and Theory Construction in Psychology." In *Scientific Concepts and Investigative Practice,* edited by Uljana Feest and Friedrich Steinle, 167–190. DE GRUYTER, October. https://doi.org/10.1515/9783110253610.167.

Fellbaum, C. 2014. "Large-Scale Lexicography in the Digital Age." *International Journal of Lexicography* 27, no. 4 (December): 378–395. https://doi.org/10.1093/ijl/ecu018.

Fenton, André A., José R. Hurtado, Jantine A.C. Broek, EunHye Park, and Bud Mishra. 2023. "Do Place Cells Dream of Deceptive Moves in a Signaling Game?" *Neuroscience* 529 (October): 129–147. https://doi.org/10.1016/j.neuroscience.2023.08.012.

Fields, Chris, and Michael Levin. 2020. “Scale-Free Biology: Integrating Evolutionary and Developmental Thinking.” *BioEssays* 42, no. 8 (August): 1900228. https://doi.org/10.1002/bies.201900228.

Figdor, Carrie. 2022. “What Could Cognition Be, If Not Human Cognition?: Individuating Cognitive Abilities in the Light of Evolution.” *Biology & Philosophy* 37, no. 6 (December). https://doi.org/10.1007/s10539-022-09880-z.

———. 2024a. “Individuating Cognitive Characters: Lessons from Praying Mantises and Plants.” *Philosophy of Science* 91, no. 4 (October): 930–949. https://doi.org/10.1017/psa.2024.10.

———. 2024b. “What Are We Talking About When We Talk About Cognition? Human, Cybernetic, and Phylogenetic Conceptual Schemes.” *JoLMA,* no. 2 (February). https://doi.org/10.30687/jolma/2723-9640/2023/02/001.

Fine, Kit. 2007. *Semantic Relationism.* 1. Malden (Mass.): Blackwell publ.

———. 2017a. “A Theory of Truthmaker Content I: Conjunction, Disjunction and Negation.” *Journal of Philosophical Logic* 46, no. 6 (December): 625–674. https://doi.org/10.1007/s10992-016-9413-y.

———. 2017b. “A Theory of Truthmaker Content II: Subject-matter, Common Content, Remainder and Ground.” *Journal of Philosophical Logic* 46, no. 6 (December): 675–702. https://doi.org/10.1007/s10992-016-9419-5.

Finger, Stanley. 2001. *Origins of Neuroscience: A History of Explorations into Brain Function.* Oxford: Oxford Univ. Press.

Fini, Chiara, Ilenia Falcinelli, and Anna M. Borghi. 2025. “Technological Concepts: Object and Tool of Knowledge Outsourcing at Different Ages.” *Topoi* (May). https://doi.org/10.1007/s11245-025-10206-3.

Firestone, Chaz, and Brian J. Scholl. 2016. “Cognition Does Not Affect Perception: Evaluating the Evidence for “Top-down” Effects.” *Behavioral and Brain Sciences* 39:e229. https://doi.org/10.1017/S0140525X15000965.

Firth, John Rupert. 1957. “A Synopsis of Linguistic Theory 1930-55.” In *Studies in Linguistic Analysis,* 1–32. Publications of the Philological Society. Oxford: Blackwell.

Flack, Jessica C. 2017. “Coarse-Graining as a Downward Causation Mechanism.” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 375, no. 2109 (December): 20160338. https://doi.org/10.1098/rsta.2016.0338.

Fleck, Ludwik. (1935) 2008. *Genesis and Development of a Scientific Fact.* Edited by Thaddeus J. Trenn and Robert K. Merton. Translated by Fred Bradley and Thaddeus J. Trenn. Chicago: Univ. of Chicago Press.

Floridi, Luciano. 2010. *Information: A Very Short Introduction.* 225. Oxford; New York: Oxford University Press.

———. 2011. *The Philosophy of Information.* Oxford ; New York: Oxford University Press.

Fodor, Jerry A. 1975. *The Language of Thought.* New York: Crowell.

———. 1992. *A Theory of Content and Other Essays.* The MIT Press. https://doi.org/10.7551/mitpress/6765.001.0001.

———. 1998. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind.* 2. Cambridge, MA: MIT Pr.

Frean, Marcus, and Edward R. Abraham. 2001. "Rock–Scissors–Paper and the Survival of the Weakest." *Proceedings of the Royal Society of London. Series B: Biological Sciences* 268, no. 1474 (July): 1323–1327. https://doi.org/10.1098/rspb.2001.1670.

Freeberg, Todd M, D L Book, Hwayoung Jung, and Steven C Kyle. 2021. "Communication, Cues, and Signals." In *Encyclopedia of Evolutionary Psychological Science,* edited by Todd K Shackelford and Viviana A Weekes-Shackelford, 1206–1214. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-19650-3_2728.

French, Steven. 2020. *There Are No Such Things As Theories.* Oxford University PressOxford, February. https://doi.org/10.1093/oso/9780198848158.001.0001.

Frick, Ramiro, Leonardo Bich, and Alvaro Moreno. 2019. "An Organisational Approach to Biological Communication." *Acta Biotheoretica* 67, no. 2 (June): 103–128. https://doi.org/10.1007/s10441-019-09342-2.

Frigg, Roman, and James Nguyen. 2020. *Modelling Nature: An Opinionated Introduction to Scientific Representation.* Vol. 427. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-45153-0.

Galantucci, Bruno. 2005. "An Experimental Study of the Emergence of Human Communication Systems." *Cognitive Science* 29, no. 5 (September): 737–767. https://doi.org/10.1207/s15516709cog0000_34.

Galen. 1968. *On the Usefulness of the Parts of the Body.* Translated by Margaret T. May. Ithaca, NY: Cornell University Press.

Gallagher, Shaun. 2005. *How the Body Shapes the Mind.* Oxford: Clarendon Press.

Gatenby, Robert A., and B. Roy Frieden. 2007. "Information Theory in Living Systems, Methods, Applications, and Challenges." *Bulletin of Mathematical Biology* 69, no. 2 (February): 635–657. https://doi.org/10.1007/s11538-006-9141-5.

Gauker, Christopher. 2015. "The Illusion of Semantic Reference." In *On Reference,* edited by Andrea Bianchi, 21–39. Oxford University Press, March. https://doi.org/10.1093/acprof:oso/9780198714088.003.0002.

Gebharter, Alexander, and Bert Leuridan. 2024. "Modelling Cyclic Causal Structures." In *The Routledge Handbook of Causality and Causal Methods,* 1st ed., 269–280. New York: Routledge, December. Accessed February 4, 2025. https://doi.org/10.4324/9781003528937-30.

Geertz, Clifford. 1973. "Thick Description: Toward an Interpretive Theory of Culture." In *The Interpretation of Cultures. Selected Essays,* 3–30. New York: Basic Books, Inc.

Gentner, Dedre, and Michael Jeziorski. 1993. "The Shift from Metaphor to Analogy in Western Science." In *Metaphor and Thought,* 2nd ed, edited by Andrew Ortony, 447–480. Cambridge (Mass.): Cambridge university press.

Gershman, Samuel J, Petra EM Balbi, C Randy Gallistel, and Jeremy Gunawardena. 2021. "Reconsidering the Evidence for Learning in Single Cells." Edited by Laura L Colgin. *eLife* 10 (January): e61907. https://doi.org/10.7554/eLife.61907.

Gibson, James J. 1950. *The Perception of the Visual World.* Boston, MA: Houghton Mifflin Company.

———. 1960. "The Concept of the Stimulus in Psychology." *American Psychologist* 15, no. 11 (November): 694–703. https://doi.org/10.1037/h0047037.

———. 1966. *The Senses Considered as Perceptual Systems.* Boston, MA: Houghton Mifflin Company.

Gibson, James J. 1986. *The Ecological Approach to Visual Perception.* Hillsdale, NJ: Erlbaum.

———. 2017. "The Theory of Affordances." In *Perceiving, Acting and Knowing: Toward an Ecological Psychology,* 0th ed., edited by Robert Shaw and John Bransford, 67–82. Routledge, March. https://doi.org/10.4324/9781315467931.

Giere, Ronald N. 1973. "History and Philosophy of Science: Intimate Relationship or Marriage of Convenience?" *The British Journal for the Philosophy of Science* 24, no. 3 (September): 282–297. https://doi.org/10.1093/bjps/24.3.282.

———. 1985. "Philosophy of Science Naturalized." *Philosophy of Science* 52 (3): 331–356.

———. 2006. *Scientific Perspectivism.* University of Chicago Press.

———. 2011. "History and Philosophy of Science: Thirty-Five Years Later." In *Integrating History and Philosophy of Science,* edited by Seymour Mauskopf and Tad Schmaltz, 263:59–65. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-1745-9_5.

Giere, Ronald N., and Barton Moffatt. 2003. "Distributed Cognition: Where the Cognitive and the Social Merge." *Social Studies of Science* 33, no. 2 (April): 301–310. https://doi.org/10.1177/03063127030332017.

Gładziejewski, Paweł. 2016. "Predictive Coding and Representationalism." *Synthese* 193, no. 2 (February): 559–582. https://doi.org/10.1007/s11229-015-0762-9.

Gładziejewski, Paweł, and Marcin Miłkowski. 2017. "Structural Representations: Causally Relevant and Different from Detectors." *Biology & Philosophy* 32, no. 3 (June): 337–355. https://doi.org/10.1007/s10539-017-9562-6.

Glaser, Joshua I., Ari S. Benjamin, Raeed H. Chowdhury, Matthew G. Perich, Lee E. Miller, and Konrad P. Kording. 2020. "Machine Learning for Neural Decoding." *eneuro* 7, no. 4 (July): ENEURO.0506–19.2020. https://doi.org/10.1523/eneuro.0506-19.2020.

Gleick, James. 2011. *The Information: A History, a Theory, a Flood.* New York: Pantheon Books.

Glennan, Stuart. 2017. *The New Mechanical Philosophy.* Vol. 1. Oxford University Press, September. https://doi.org/10.1093/oso/9780198779711.001.0001.

Główka, Krzysztof, Julian Zubek, and Joanna Rączaszek-Leonardi. 2024. "Context-Dependent Communication under Environmental Constraints." *Cognitive Systems Research* 88 (December): 101293. https://doi.org/10.1016/j.cogsys.2024.101293.

Godfrey-Smith, Peter. 2000. "Information, Arbitrariness, and Selection: Comments on Maynard Smith." *Philosophy of Science* 67 (2): 202–207.

———. 2007. "Information in Biology." In *The Cambridge Companion to the Philosophy of Biology,* edited by David L. Hull and Michael Ruse, 103–119. Cambridge Companions to Philosophers. New York: Cambridge University Press.

———. 2011. "Senders, Receivers, and Genetic Information: Comments on Bergstrom and Rosvall." *Biology & Philosophy* 26, no. 2 (March): 177–181. https://doi.org/10.1007/s10539-010-9206-6.

———. 2012. "Signals, Icons, and Beliefs." In *Millikan and Her Critics,* 1st ed., edited by Dan Ryder, Justine Kingsbury, and Kenneth Williford, 41–62. Wiley, November. https://doi.org/10.1002/9781118328118.ch2.

———. 2013. "Information and Influence in Sender-Receiver Models, with Applications to Animal Behavior." In *Animal Communication Theory: Information and Influence,* edited by Ulrich E. Stegmann, 377–396. Cambridge ; New York: Cambridge University Press.

———. 2014. "Sender-Receiver Systems within and between Organisms." *Philosophy of Science* 81, no. 5 (December): 866–878. https://doi.org/10.1086/677686.

———. 2016. "Mind, Matter, and Metabolism:" *Journal of Philosophy* 113 (10): 481–506. https://doi.org/10.5840/jphil20161131034.

———. 2018. *Other Minds: The Octopus and the Evolution of Intelligent Life.* London: William Collins.

Gold, Ian, and Daniel Stoljar. 1999. "A Neuron Doctrine in the Philosophy of Neuroscience." *Behavioral and Brain Sciences* 22, no. 5 (October): 809–830. https://doi.org/10.1017/S0140525X99002198.

Good, I. J. 1979. "Studies in the History of Probability and Statistics. XXXVII A. M. Turing's Statistical Work in World War II." *Biometrika* 66 (2): 393–396. https://doi.org/10.1093/biomet/66.2.393.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. 2016. *Deep Learning.* Cambridge, Massachusetts: The MIT Press.

Gould, Stephen J., and Richard C. Lewontin. 1979. "The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme." *Proceedings of the Royal Society of London. Series B. Biological Sciences* 205, no. 1161 (September): 581–598. https://doi.org/10.1098/rspb.1979.0086.

Gould, Stephen Jay. 1977. *Ontogeny and Phylogeny.* Cambridge, Mass: Belknap Press of Harvard University Press.

Green, Mitchell S. 2007. *Self-Expression.* Oxford: Oxford University Press.

Grice, H. Paul. 1957. "Meaning." *The Philosophical Review* 66, no. 3 (July): 377. https://doi.org/10.2307/2182440.

———. 1969. "Utterer's Meaning and Intention." *The Philosophical Review* 78, no. 2 (April): 147. https://doi.org/10.2307/2184179.

———. 1995. *Studies in the Way of Words.* Cambridge, Mass.: Harvard Univ. Pr.

Griesemer, James R. 2005. "The Informational Gene and the Substantial Body: On the Generalization of Evolutionary Theory by Abstraction." In *Idealization XII: Correcting the Model: Idealization and Abstraction in the Sciences,* edited by Martin R. Jones and Nancy Cartwright. BRILL, January. https://doi.org/10.1163/9789401202732.

Griffiths, P. E., and R. D. Gray. 1994. "Developmental Systems and Evolutionary Explanation." Edited by John Smylie. *Journal of Philosophy* 91 (6): 277–304. https://doi.org/10.2307/2940982.

Griffiths, Paul E. 1993. "Functional Analysis and Proper Functions." *The British Journal for the Philosophy of Science* 44, no. 3 (September): 409–422. https://doi.org/10.1093/bjps/44.3.409.

———. 2001. "Genetic Information: A Metaphor in Search of a Theory." *Philosophy of Science* 68 (3): 394–412.

———. 2016. "Proximate and Ultimate Information in Biology." In *The Philosophy of Philip Kitcher,* edited by Mark B. Couch and Jessica Pfeifer, 74–97. Oxford University Press, July. https://doi.org/10.1093/acprof:oso/9780199381357.003.0004.

Grootendorst, Maarten. 2020. *KeyBERT: Minimal Keyword Extraction with BERT.* Zenodo. V. v0.1.3. https://doi.org/10.5281/zenodo.4461265.

Grootendorst, Maarten. 2022. "BERTopic: Neural Topic Modeling with a Class-Based TF-IDF Procedure." *arXiv preprint arXiv:2203.05794.*

Grush, Rick. 2004. "The Emulation Theory of Representation: Motor Control, Imagery, and Perception." *The Behavioral and Brain Sciences* 27, no. 3 (June): 377–396, discussion 396–442.

Guha, R., Rob McCool, and Eric Miller. 2003. "Semantic Search." In *Proceedings of the Twelfth International Conference on World Wide Web - WWW '03,* 700. Budapest, Hungary: ACM Press. https://doi.org/10.1145/775152.775250.

Gunka, Katrin, and Fabian M. Commichau. 2012. "Control of Glutamate Homeostasis in *Bacillus Subtilis* : A Complex Interplay between Ammonium Assimilation, Glutamate Biosynthesis and Degradation." *Molecular Microbiology* 85, no. 2 (July): 213–224. https://doi.org/10.1111/j.1365-2958.2012.08105.x.

Haak, Koen V., and Christian F. Beckmann. 2020. "Understanding Brain Organisation in the Face of Functional Heterogeneity and Functional Multiplicity." *NeuroImage* 220 (October): 117061. https://doi.org/10.1016/j.neuroimage.2020.117061.

Hacking, Ian. (1983) 2007. *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science.* Cambridge: Cambridge Univ. Press. https://doi.org/10.1017/CBO9780511814563.

Hammond, Adam. 2017. "The Double Bind of Validation: Distant Reading and the Digital Humanities' "Trough of Disillusionment"." *Literature Compass* 14, no. 8 (August): e12402. https://doi.org/10.1111/lic3.12402.

Han, Kyung-Seok, Chong Guo, Christopher H. Chen, Laurens Witter, Tomas Osorno, and Wade G. Regehr. 2018. "Ephaptic Coupling Promotes Synchronous Firing of Cerebellar Purkinje Cells." *Neuron* 100, no. 3 (November): 564–578.e3. https://doi.org/10.1016/j.neuron.2018.09.018.

Han, Lei. 2021. "The Idealistic Elements in Modern Semiotic Studies: With Particular Recourse to the Umwelt Theory." *Concentric Literary and Cultural Studies* 47, no. 1 (March). https://doi.org/10.6240/concentric.lit.202103_47(1).0006.

Hanson, Alison. 2021. "Spontaneous Electrical Low-Frequency Oscillations: A Possible Role in *Hydra* and All Living Systems." *Philosophical Transactions of the Royal Society B: Biological Sciences* 376, no. 1820 (March): 20190763. https://doi.org/10.1098/rstb.2019.0763.

Haraway, Donna. 1992. *Primate Visions: Gender, Race, and Nature in the World of Modern Science.* London: Verso.

Hardcastle, Valerie Gray. 1996. *How to Build a Theory in Cognitive Science.* Albany: State University of New York Press.

Harnad, Stevan. 1990. "The Symbol Grounding Problem." *Physica D: Nonlinear Phenomena* 42, nos. 1-3 (June): 335–346. https://doi.org/10.1016/0167-2789(90)90087-6.

Harper, L. J., K. Schönrogge, K. Y. Lim, P. Francis, and C. P. Lichtenstein. 2004. "Cynipid Galls: Insect-induced Modifications of Plant Development Create Novel Plant Organs." *Plant, Cell & Environment* 27, no. 3 (February): 327–335. https://doi.org/10.1046/j.1365-3040.2004.01145.x.

Harris, Zellig S. 1981. "Distributional Structure." In *Papers on Syntax,* edited by Henry Hiż, 3–22. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-009-8467-7_1.

Harrison, David. 2023. "Mind the Matter: Active Matter, Basal Cognition, and the Making of Bio-Inspired Artificial Intelligence." PhD diss., University of Cambridge, July. Accessed December 23, 2024.

Harrison, David, Wiktor Rorot, and Urte Laukaityte. 2022. "Mind the Matter: Active Matter, Soft Robotics, and the Making of Bio-Inspired Artificial Intelligence." *Frontiers in Neurorobotics* 16 (December): 880724. https://doi.org/10.3389/fnbot.2022.880724.

Haueis, Philipp. 2021. "The Death of the Cortical Column? Patchwork Structure and Conceptual Retirement in Neuroscientific Practice." *Studies in History and Philosophy of Science Part A* 85 (February): 101–113. https://doi.org/10.1016/j.shpsa.2020.09.010.

———. 2024. "A Generalized Patchwork Approach to Scientific Concepts." *The British Journal for the Philosophy of Science* 75, no. 3 (September): 741–768. https://doi.org/10.1086/716179.

Haugeland, John. 1993. *Artificial Intelligence: The Very Idea.* Cambridge, Mass.: MIT Press.

Hauser, Marc D. 2000. *The Evolution of Communication.* Cambridge, Mass.: MIT Press.

Hawking, Stephen W., and Leonard Mlodinow. 2010. *The Grand Design.* New York, NY: Bantam Books.

Haxby, James V., Andrew C. Connolly, and J. Swaroop Guntupalli. 2014. "Decoding Neural Representational Spaces Using Multivariate Pattern Analysis." *Annual Review of Neuroscience* 37, no. 1 (July): 435–456. https://doi.org/10.1146/annurev-neuro-062012-170325.

Haxby, James V., M. Ida Gobbini, Maura L. Furey, Alumit Ishai, Jennifer L. Schouten, and Pietro Pietrini. 2001. "Distributed and Overlapping Representations of Faces and Objects in Ventral Temporal Cortex." *Science* 293, no. 5539 (September): 2425–2430. https://doi.org/10.1126/science.1063736.

Heft, Harry. 2001. *Ecological Psychology in Context.* Psychology Press, July. https://doi.org/10.4324/9781410600479.

Heintz, Christophe, and Thom Scott-Phillips. 2023. "Expression Unleashed: The Evolutionary and Cognitive Foundations of Human Communication." *Behavioral and Brain Sciences* 46:e1. https://doi.org/10.1017/S0140525X22000012.

Hempel, Carl G., and Paul Oppenheim. 1948. "Studies in the Logic of Explanation." *Philosophy of Science* 15, no. 2 (April): 135–175. https://doi.org/10.1086/286983.

Herron, Matthew D. 2016. "Origins of Multicellular Complexity: Volvox and the Volvocine Algae." *Molecular Ecology* 25, no. 6 (March): 1213–1223. https://doi.org/10.1111/mec.13551.

Hesse, Mary. 1966. *Models and Analogies in Science.* Notre Dame: University of Notre Dame Press.

———. 2017. "Models and Analogies." In *A Companion to the Philosophy of Science,* 1st ed., edited by W. H. Newton-Smith, 299–307. Wiley, September. https://doi.org/10.1002/9781405164481.ch44.

Heuser, Ryan, and Long Le-Khac. 2012. *A Quantitative Literary History of 2,958 Nineteenth-Century British Novels : The Semantic Cohort Method.* 4. Stanford: Stanford Literary Lab.

Hinrichs, Uta, Stefania Forlini, and Bridget Moynihan. 2019. "In Defense of Sandcastles: Research Thinking through Visualization in Digital Humanities." *Digital Scholarship in the Humanities* 34, no. Supplement_1 (December): i80–i99. https://doi.org/10.1093/llc/fqy051.

Hodges, Bert H., and Reuben M. Baron. 1992. "Values as Constraints on Affordances: Perceiving and Acting Properly." *Journal for the Theory of Social Behaviour* 22, no. 3 (September): 263–294. https://doi.org/10.1111/j.1468-5914.1992.tb00220.x.

Hodges, Bert H., and Joanna Rączaszek-Leonardi. 2021. "Ecological Values Theory: Beyond Conformity, Goal-Seeking, and Rule-Following in Action and Interaction." *Review of General Psychology* (September): 108926802110481. https://doi.org/10.1177/10892680211048174.

Hoeber, Orland, Morgan Harvey, Milad Momeni, Abbas Pirmoradi, and David Gleeson. 2024. "Exploratory Search in Digital Humanities: A Study of Visual Keyword/Result Linking." *Proceedings of the Association for Information Science and Technology* 61, no. 1 (October): 161–171. https://doi.org/10.1002/pra2.1017.

Hoffmeyer, Jesper. 2008. *Biosemiotics: An Examination into the Signs of Life and the Life of Signs.* Edited by Donald Favareau. Translated by Jesper Hoffmeyer and Donald Favareau. 2. Scranton, Pa.: University of Scranton Press.

Hohwy, Jakob. 2013. *The Predictive Mind.* Oxford, United Kingdom ; New York, NY, United States of America: Oxford University Press.

Holland, John H. 2010. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence.* Cambridge, Mass.: MIT Press.

———. 2012. *Signals and Boundaries: Building Blocks for Complex Adaptive Systems.* Cambridge, Mass: MIT Press.

Horák, Aleš, and Adam Rambousek. 2018. "Lexicography and Natural Language Processing." In *The Routledge Handbook of Lexicography,* edited by Pedro Antonio Fuertes Olivera. London ; New York: Routledge.

Hubel, David H., and Torsten N. Wiesel. 1977. "Ferrier Lecture - Functional Architecture of Macaque Monkey Visual Cortex." *Proceedings of the Royal Society of London. Series B. Biological Sciences* 198, no. 1130 (July): 1–59. https://doi.org/10.1098/rspb.1977.0085.

Hubert, Gilles. 2024. "Bad Practices in Open Peer Review: Lessons Learned from Mining MDPI Open Reports." In *IX Annual Conference of the Italian Association for the Promotion of Open Science.* Pisa: Zenodo, November. https://doi.org/10.5281/ZENODO.14098794.

Hughes, Brian. 2018. *Psychology in Crisis.* London: Palgrave McMillan.

Hulbert, A.J, and Paul Lewis Else. 1999. "Membranes as Possible Pacemakers of Metabolism." *Journal of Theoretical Biology* 199, no. 3 (August): 257–274. https://doi.org/10.1006/jtbi.1999.0955.

Hummert, Sabine, Katrin Bohl, David Basanta, Andreas Deutsch, Sarah Werner, Günter Theißen, Anja Schroeter, and Stefan Schuster. 2014. "Evolutionary Game Theory: Cells as Players." *Molecular BioSystems* 10 (12): 3044–3065. https://doi.org/10.1039/C3MB70602H.

Humphreys, Paul. 2002. "Computational Models." *Philosophy of Science* 69, no. S3 (September): S1–S11. https://doi.org/10.1086/341763.

———. 2019. "Knowledge Transfer across Scientific Disciplines." *Studies in History and Philosophy of Science Part A* 77 (October): 112–119. https://doi.org/10.1016/j.shpsa.2017.11.001.

Hutchins, Edwin. 2006. *Cognition in the Wild.* Cambridge, Mass.: MIT Press.

Huttegger, Simon M. 2007. "Robustness in Signaling Games." *Philosophy of Science* 74, no. 5 (December): 839–847. https://doi.org/10.1086/525625.

Hutter, Marcus. 2007. "Algorithmic Information Theory." *Scholarpedia* 2 (3): 2519. https://doi.org/10.4249/scholarpedia.2519.

Ikram, Muhammad Touseef, and Muhammad Tanvir Afzal. 2019. "Aspect Based Citation Sentiment Analysis Using Linguistic Patterns for Better Comprehension of Scientific Knowledge." *Scientometrics* 119, no. 1 (April): 73–95. https://doi.org/10.1007/s11192-019-03028-9.

Incitti, Francesca, Federico Urli, and Lauro Snidaro. 2023. "Beyond Word Embeddings: A Survey." *Information Fusion* 89 (January): 418–436. https://doi.org/10.1016/j.inffus.2022.08.024.

Ioannidis, John P. A. 2005. "Why Most Published Research Findings Are False." *PLoS Medicine* 2, no. 8 (August): e124. https://doi.org/10.1371/journal.pmed.0020124.

Irvine, Elizabeth. 2021. "The Role of Replication Studies in Theory Building." *Perspectives on Psychological Science* 16, no. 4 (July): 844–853. https://doi.org/10.1177/1745691620970558.

Isaac, Alistair M C. 2019. "The Semantics Latent in Shannon Information." *The British Journal for the Philosophy of Science* 70, no. 1 (March): 103–125. https://doi.org/10.1093/bjps/axx029.

Jabareen, Yosef. 2009. "Building a Conceptual Framework: Philosophy, Definitions, and Procedure." *International Journal of Qualitative Methods* 8, no. 4 (December): 49–62. https://doi.org/10.1177/160940690900800406.

Jablonka, Eva. 2002. "Information: Its Interpretation, Its Inheritance, and Its Sharing." *Philosophy of Science* 69, no. 4 (December): 578–605. https://doi.org/10.1086/344621.

Jadacki, Jacek J. 1990. *O rozumieniu. Z filozoficznych podstaw semiotyki.* Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.

Jakobson, Roman. (1960) 1981. "Linguistics and Poetics." In *Poetry of Grammar and Grammar of Poetry,* 18–51. De Gruyter Mouton. Accessed January 17, 2025. https://doi.org/10.1515/9783110802122.18.

———. 1980. "Sign and System of Language: A Reassessment of Saussure's Doctrine." Translated by B. Hrushovski. *Poetics Today* 2 (1a): 33. https://doi.org/10.2307/1772350.

James, William. 1981. *The Principles of Psychology.* Cambridge, Mass: Harvard University Press.

Jarnicki, Paweł, and Hajo Greif. 2024. "The 'Aristotle Experience' Revisited: Thomas Kuhn Meets Ludwik Fleck on the Road to *Structure*." *Archiv für Geschichte der Philosophie* 106, no. 2 (June): 313–349. https://doi.org/10.1515/agph-2020-0160.

Jiang, Xiaorui, and Jingqiang Chen. 2023. "Contextualised Segment-Wise Citation Function Classification." *Scientometrics* 128, no. 9 (September): 5117–5158. https://doi.org/10.1007/s11192-023-04778-3.

Johnson, Mark. 2010. *The Meaning of the Body: Aesthetics of Human Understanding.* Chicago, Ill.: Univ. of Chicago Press.

Johnson, R. Burke, Anthony J. Onwuegbuzie, and Lisa A. Turner. 2007. "Toward a Definition of Mixed Methods Research." *Journal of Mixed Methods Research* 1, no. 2 (April): 112–133. https://doi.org/10.1177/1558689806298224.

Kaiser, Marie I. 2019. "Normativity in the Philosophy of Science." *Metaphilosophy* 50, nos. 1-2 (January): 36–62. https://doi.org/10.1111/meta.12348.

Kalkman, David. 2019. "New Problems for Defining Animal Communication in Informational Terms." *Synthese* 196, no. 8 (August): 3319–3336. https://doi.org/10.1007/s11229-017-1598-2.

Kalman, R. E. 1960. “A New Approach to Linear Filtering and Prediction Problems.” *Journal of Basic Engineering* 82, no. 1 (March): 35–45. https://doi.org/10.1115/1.3662552.

Kant, Immanuel. (1781) 1998. *Critique of Pure Reason.* Edited by Paul Guyer and Allen W. Wood. Cambridge ; New York: Cambridge University Press.

Karlebach, Guy, and Ron Shamir. 2008. “Modelling and Analysis of Gene Regulatory Networks.” *Nature Reviews Molecular Cell Biology* 9, no. 10 (October): 770–780. https://doi.org/10.1038/nrm2503.

Kauffman, Stuart A. 2019. *A World beyond Physics: The Emergence and Evolution of Life.* New York, NY: Oxford University Press.

Keller, Evelyn Fox. 2000. “Models of and Models for: Theory and Practice in Contemporary Biology.” *Philosophy of Science* 67:S72–S86.

Kempson, Ruth M. 1990. “The Relation between Language, Mind, and Reality.” In *Mental Representations: The Interface between Language and Reality,* Reprint, edited by Ruth M. Kempson, 1–25. Cambridge: Cambridge Univ. Press.

Khalidi, Muhammad Ali. 2017. “Crosscutting Psycho-Neural Taxonomies: The Case of Episodic Memory.” *Philosophical Explorations* 20, no. 2 (May): 191–208. https://doi.org/10.1080/13869795.2017.1312501.

———. 2022. *Cognitive Ontology: Taxonomic Practices in the Mind-Brain Sciences.* New York, NY: Cambridge University Press.

Khalifa, Kareem. 2017. *Understanding, Explanation, and Scientific Knowledge.* Cambridge, United Kingdom: Cambridge University Press.

Khelfaoui, Mahdi, Yves Gingras, Maël Lemoine, and Thomas Pradeu. 2021. “The Visibility of Philosophy of Science in the Sciences, 1980–2018.” *Synthese* 199, nos. 3-4 (December): 6219–6249. https://doi.org/10.1007/s11229-021-03067-x.

Kight, Caitlin R., John M. McNamara, David W. Stephens, and Sasha R. X. Dall. 2013. “Communication as Information Use: Insights from Statistical Decision Theory.” In *Animal Communication Theory: Information and Influence,* edited by Ulrich E. Stegmann, 89–111. Cambridge ; New York: Cambridge University Press.

Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý, and Vít Suchomel. 2014. “The Sketch Engine: Ten Years On.” *Lexicography* 1, no. 1 (July): 7–36. https://doi.org/10.1007/s40607-014-0009-9.

Kim, Yoon, Yi-I. Chiu, Kentaro Hanaki, Darshan Hegde, and Slav Petrov. 2014. *Temporal Analysis of Language through Neural Language Models,* arXiv:1405.3515, May. Accessed January 27, 2023.

Kincheloe, Joe L. 2011. “Critical Ontology and Indigenous Ways of Being: Forging a Postcolonial Curriculum.” In *Key Works in Critical Pedagogy,* edited by Kecia Hayes, Shirley R. Steinberg, and Kenneth Tobin, 333–349. Rotterdam: SensePublishers. https://doi.org/10.1007/978-94-6091-397-6_25.

Kinney, Rodney, Chloe Anastasiades, Russell Authur, Iz Beltagy, Jonathan Bragg, Alexandra Buraczynski, Isabel Cachola, et al. 2025. *The Semantic Scholar Open Data Platform,* 2301.10140, April. Accessed June 20, 2025. https://doi.org/10.48550/arXiv.2301.10140.

Kirsch, Adam. 2014. “Technology Is Taking Over English Departments.” *The New Republic* (May).

Kitcher, Philip. 2003. "Battling the Undead: How (and How Not) to Resist Genetic Determinism (2000)." In *In Mendel's Mirror,* 283–300. Oxford University PressNew York, NY, March. Accessed May 11, 2025. https://doi.org/10.1093/oso/9780195151787.003.0013.

Knobe, Joshua. 2007. "Experimental Philosophy." *Philosophy Compass* 2, no. 1 (January): 81–92. https://doi.org/10.1111/j.1747-9991.2006.00050.x.

Knorr-Cetina, Karin D. 1981. "Social and Scientific Method or What Do We Make of the Distinction Between the Natural and the Social Sciences?" *Philosophy of the Social Sciences* 11, no. 3 (September): 335–359. https://doi.org/10.1177/004839318101100304.

Kolchinsky, Artemy, and David H. Wolpert. 2018. "Semantic Information, Autonomous Agency and Non-Equilibrium Statistical Physics." *Interface Focus* 8, no. 6 (December): 20180041. https://doi.org/10.1098/rsfs.2018.0041.

Komorowska-Mach, Joanna. 2019. "Introspection — One or More? Pluralism about Self-Knowledge." *Filozofia Nauki* 27, no. 1 (March). https://doi.org/10.14394/filnau.2019.0001.

Kopan, Raphael. 2012. "Notch Signaling." *Cold Spring Harbor Perspectives in Biology* 4, no. 10 (October): a011213. https://doi.org/10.1101/cshperspect.a011213.

Korbak, Tomasz. 2023. "Self-Organisation, (M, R)–Systems and Enactive Cognitive Science." *Adaptive Behavior* 31, no. 1 (February): 35–49. https://doi.org/10.1177/10597123211066155.

Korbak, Tomasz, Julian Zubek, Łukasz Kuciński, Piotr Miłoś, and Joanna Rączaszek-Leonardi. 2019. "Developmentally Motivated Emergence of Compositional Communication via Template Transfer."

———. 2021. "Interaction History as a Source of Compositionality in Emergent Communication." *Interaction Studies. Social Behaviour and Communication in Biological and Artificial Systems* 22, no. 2 (December): 212–243. https://doi.org/10.1075/is.21020.kor.

Kottur, Satwik, José M. F. Moura, Stefan Lee, and Dhruv Batra. 2017. "Natural Language Does Not Emerge 'Naturally' in Multi-Agent Dialog." *arXiv:1706.08502 [cs]* (August).

Kreer, J. 1957. "A Question of Terminology." *IRE Transactions on Information Theory* 3, no. 3 (September): 208–208. https://doi.org/10.1109/TIT.1957.1057418.

Krickel, Beate, and Mariel Goddu. 2025. "Cognitive Ontology in Terms of Cognitive Homology: The Role of Brain, Behavior, and Environment for Individuating Cognitive Categories." In *Neurocognitive Foundations of Mind,* edited by Gualtiero Piccinini. Abingdon, Oxon: Routledge.

Kriegeskorte, Nikolaus. 2008. "Representational Similarity Analysis – Connecting the Branches of Systems Neuroscience." *Frontiers in Systems Neuroscience,* https://doi.org/10.3389/neuro.06.004.2008.

Kriegeskorte, Nikolaus, and Rogier A. Kievit. 2013. "Representational Geometry: Integrating Cognition, Computation, and the Brain." *Trends in Cognitive Sciences* 17, no. 8 (August): 401–412. https://doi.org/10.1016/j.tics.2013.06.007.

Kuciński, Łukasz, Tomasz Korbak, Paweł Kołodziej, and Piotr Miłoś. 2021. "Catalytic Role Of Noise And Necessity Of Inductive Biases In The Emergence Of Compositional Communication." *arXiv:2111.06464 [cs]* (November).

Kuhn, Thomas S. (1962) 2012. *The Structure of Scientific Revolutions.* Edited by Ian Hacking. Chicago ; London: The University of Chicago Press.

Kulkarni, Vivek, Rami Al-Rfou, Bryan Perozzi, and Steven Skiena. 2015. "Statistically Significant Detection of Linguistic Change." In *Proceedings of the 24th International Conference on World Wide Web,* 625–635. Florence Italy: International World Wide Web Conferences Steering Committee, May. https://doi.org/10.1145/2736277.2741627.

Kull, Kalevi. 2000. "Trends in Theoretical Biology: The 20th Century." *Aquinas* 43 (2): 235–249.

Kurien, B, and R Scofield. 2006. "Western Blotting." *Methods* 38, no. 4 (April): 283–293. https://doi.org/10.1016/j.ymeth.2005.11.007.

Lakoff, George, and Mark Johnson. 1980. *Metaphors We Live By.* Chicago: University of Chicago Press.

Landauer, Rolf. 1961. "Irreversibility and Heat Generation in the Computing Process." *IBM Journal of Research and Development* 5, no. 3 (July): 183–191. https://doi.org/10.1147/rd.53.0183.

Landauer, Thomas K., and Susan T. Dumais. 1997. "A Solution to Plato's Problem: The Latent Semantic Analysis Theory of Acquisition, Induction, and Representation of Knowledge." *Psychological Review* 104, no. 2 (April): 211–240. https://doi.org/10.1037/0033-295X.104.2.211.

Latour, Bruno. 2005. *Reassembling the Social: An Introduction to Actor-Network-Theory.* Oxford ; New York: Oxford University Press.

Latour, Bruno, and Stève Woolgar. 1979. *Laboratory Life: The Social Construction of Scientific Facts.* vol. 80. Beverly Hills London: Sage Publications.

Lawley, Grace O., Peter A. Heeman, Jill K. Dolata, Eric Fombonne, and Steven Bedrick. 2023. "A Statistical Approach for Quantifying Group Difference in Topic Distributions Using Clinical Discourse Samples." In *Proceedings of the 24th Annual Meeting of the Special Interest Group on Discourse and Dialogue,* edited by Svetlana Stoyanchev, Shafiq Joty, David Schlangen, Ondrej Dusek, Casey Kennington, and Malihe Alikhani, 55–65. Prague, Czechia: Association for Computational Linguistics, September. https://doi.org/10.18653/v1/2023.sigdial-1.5.

Laymon, Ronald. 1991. "Thought Experiments by Stevin, Mach and Gouy: Thought Experiments as Ideal Limits and as Semantic Domains." In *Thought Experiments in Science and Philosophy,* edited by Tamara Horowitz and Gerald J. Massey. Center for Philosophy of Science Publications in Philosophy of Science. Savage, Md: Rowman & Littlefield.

Lazaridou, Angeliki, Karl Moritz Hermann, Karl Tuyls, and Stephen Clark. 2018. "Emergence of Linguistic Communication from Referential Games with Symbolic and Pixel Input." *arXiv:1804.03984 [cs]* (April).

Lean, Oliver M. 2014. "Getting the Most out of Shannon Information." *Biology & Philosophy* 29, no. 3 (May): 395–413. https://doi.org/10.1007/s10539-013-9410-2.

Lean, Oliver M., Luca Rivelli, and Charles H. Pence. 2023. "Digital Literature Analysis for Empirical Philosophy of Science." *The British Journal for the Philosophy of Science* 74, no. 4 (December): 875–898. https://doi.org/10.1086/715049.

Lefèvre, Christopher T., and Dennis A. Bazylinski. 2013. "Ecology, Diversity, and Evolution of Magnetotactic Bacteria." *Microbiology and Molecular Biology Reviews* 77, no. 3 (September): 497–526. https://doi.org/10.1128/MMBR.00021-13.

Lenci, Alessandro. 2008. "Distributional Semantics in Linguistic and Cognitive Research." *Italian Journal of Linguistics* 20 (1): 1–31.

———. 2018. "Distributional Models of Word Meaning." *Annual Review of Linguistics* 4, no. 1 (January): 151–171. https://doi.org/10.1146/annurev-linguistics-030514-125254.

Levenstein, Daniel, Veronica A. Alvarez, Asohan Amarasingham, Habiba Azab, Zhe S. Chen, Richard C. Gerkin, Andrea Hasenstaub, et al. 2023. "On the Role of Theory and Modeling in Neuroscience." *The Journal of Neuroscience* 43, no. 7 (February): 1074–1088. https://doi.org/10.1523/JNEUROSCI.1179-22.2022.

Levin, Janet. 2021. "Functionalism." In *The Stanford Encyclopedia of Philosophy,* Winter 2021, edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University.

Levin, Michael. 2019. "The Computational Boundary of a "Self": Developmental Bioelectricity Drives Multicellularity and Scale-Free Cognition." *Frontiers in Psychology* 10:2688. https://doi.org/10.3389/fpsyg.2019.02688.

———. 2021. "Bioelectric Signaling: Reprogrammable Circuits Underlying Embryogenesis, Regeneration, and Cancer." *Cell* 184, no. 8 (April): 1971–1989. https://doi.org/10.1016/j.cell.2021.02.034.

———. 2022. "Technological Approach to Mind Everywhere: An Experimentally-Grounded Framework for Understanding Diverse Bodies and Minds." *Frontiers in Systems Neuroscience* 16 (March): 768201. https://doi.org/10.3389/fnsys.2022.768201.

———. 2023. "Bioelectric Networks: The Cognitive Glue Enabling Evolutionary Scaling from Physiology to Mind." *Animal Cognition* 26, no. 6 (November): 1865–1891. https://doi.org/10.1007/s10071-023-01780-3.

Levin, Michael, and Daniel C. Dennett. 2020. "Cognition All the Way Down." *Aeon* (October).

Levy, Arnon. 2011. "Information in Biology: A Fictionalist Account." *Noûs* 45 (4): 640–657.

Lewens, Tim. 2014. "KIM STERELNY *The Evolved Apprentice: How Evolution Made Humans Unique.*" *The British Journal for the Philosophy of Science* 65, no. 1 (March): 185–189. https://doi.org/10.1093/bjps/axt003.

Lewis, David K. 1969. *Convention: A Philosophical Study.* Cambridge, Massachusetts: Wiley-Blackwell.

Li, Ming, and Paul Vitányi. 1993. *An Introduction to Kolmogorov Complexity and Its Applications.* New York; Berlin; Heidelberg: Springer Verlag.

Li, Siqiao, Tadashi Dohi, and Hiroyuki Okamura. 2023. "Nonhomogeneous Markov Process Modeling for Software Reliability Assessment." *IEEE Transactions on Reliability* 72, no. 4 (December): 1540–1555. https://doi.org/10.1109/tr.2023.3316891.

Lindsay, Grace W. 2021. "Convolutional Neural Networks as a Model of the Visual System: Past, Present, and Future." *Journal of Cognitive Neuroscience* 33, no. 10 (September): 2017–2031. https://doi.org/10.1162/jocn_a_01544.

Liu, Jintao, Arthur Prindle, Jacqueline Humphries, Marçal Gabalda-Sagarra, Munehiro Asally, Dong-yeon D. Lee, San Ly, Jordi Garcia-Ojalvo, and Gürol M. Süel. 2015. "Metabolic Co-Dependence Gives Rise to Collective Oscillations within Biofilms." *Nature* 523, no. 7562 (July): 550–554. https://doi.org/10.1038/nature14660.

Lo, Kyle, Lucy Lu Wang, Mark Neumann, Rodney Kinney, and Dan S. Weld. 2020. *S2ORC: The Semantic Scholar Open Research Corpus,* arXiv:1911.02782, July. Accessed August 30, 2024.

Longino, Helen E. 1990. *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry.* Princeton, N.J: Princeton University Press.

———. 1995. "Gender, Politics, and the Theoretical Virtues." *Synthese* 104, no. 3 (September): 383–397. https://doi.org/10.1007/BF01064506.

Lotka, Alfred J. 1910. “Contribution to the Theory of Periodic Reactions.” *The Journal of Physical Chemistry* 14, no. 3 (March): 271–274. https://doi.org/10.1021/j150111a004.

———. 1920. “Analytical Note on Certain Rhythmic Relations in Organic Systems.” *Proceedings of the National Academy of Sciences of the United States of America* 6, no. 7 (July): 410–415. https://doi.org/10.1073/pnas.6.7.410.

Love, Alan C. 2013. “Theory Is as Theory Does: Scientific Practice and Theory Structure in Biology.” *Biological Theory* 7, no. 4 (June): 325–337. https://doi.org/10.1007/s13752-012-0046-2.

Lu, Chao, Ying Ding, and Chengzhi Zhang. 2017. “Understanding the Impact Change of a Highly Cited Article: A Content-Based Citation Analysis.” *Scientometrics* 112, no. 2 (August): 927–945. https://doi.org/10.1007/s11192-017-2398-7.

Lund, Kevin, and Curt Burgess. 1996. “Producing High-Dimensional Semantic Spaces from Lexical Co-Occurrence.” *Behavior Research Methods, Instruments, & Computers* 28, no. 2 (June): 203–208. https://doi.org/10.3758/BF03204766.

Lyon, Pamela. 2006. “The Biogenic Approach to Cognition.” *Cognitive Processing* 7, no. 1 (March): 11–29. https://doi.org/10.1007/s10339-005-0016-8.

———. 2015. “The Cognitive Cell: Bacterial Behavior Reconsidered.” *Frontiers in Microbiology* 6.

———. 2020. “Of What Is “Minimal Cognition” the Half-Baked Version?” *Adaptive Behavior* 28, no. 6 (December): 407–424. https://doi.org/10.1177/1059712319871360.

Lyon, Pamela, Fred Keijzer, Detlev Arendt, and Michael Levin. 2021. “Reframing Cognition: Getting down to Biological Basics.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 376, no. 1820 (March): 20190750. https://doi.org/10.1098/rstb.2019.0750.

Lyon, Pamela, and Franz Kuchling. 2021. “Valuing What Happens: A Biogenic Approach to Valence and (Potentially) Affect.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 376, no. 1820 (March): 20190752. https://doi.org/10.1098/rstb.2019.0752.

Machamer, Peter, Lindley Darden, and Carl F. Craver. 2000. “Thinking about Mechanisms.” *Philosophy of Science* 67 (1): 1–25. https://doi.org/10.1086/392759.

Machamer, Peter, and Alison Springle. 2016. “Perceiving and Knowing as Activities.” In *Philosophy as the Foundation of Knownledge Action Ethos.* Wydawnictwo UŁ. Accessed November 26, 2025. https://doi.org/10.18778/8088-538-7.06.

MacKay, Donald MacCrimmon. 1969. *Information, Mechanism and Meaning.* Cambridge, MA; London: The MIT Press.

Maclaurin, James. 2011. “Commentary on “The Transmission Sense of Information” by Carl T. Bergstrom and Martin Rosvall.” *Biology & Philosophy* 26, no. 2 (March): 191–194. https://doi.org/10.1007/s10539-010-9233-3.

Malaterre, Christophe, Jean-François Chartier, and Francis Lareau. 2020. “The Recipes of Philosophy of Science: Characterizing the Semantic Structure of Corpora by Means of Topic Associative Rules.” Edited by Thilo Gross. *PLOS ONE* 15, no. 11 (November): e0242353. https://doi.org/10.1371/journal.pone.0242353.

Malaterre, Christophe, Jean-François Chartier, and Davide Pulizzotto. 2019. “What Is This Thing Called *Philosophy of Science* ? A Computational Topic-Modeling Perspective, 1934–2015.” *HOPOS: The Journal of the International Society for the History of Philosophy of Science* 9, no. 2 (September): 215–249. https://doi.org/10.1086/704372.

Malaterre, Christophe, and Francis Lareau. 2022. "The Early Days of Contemporary Philosophy of Science: Novel Insights from Machine Translation and Topic-Modeling of Non-Parallel Multilingual Corpora." *Synthese* 200, no. 3 (June): 242. https://doi.org/10.1007/s11229-022-03722-x.

———. 2023. "The Emergence of Astrobiology: A Topic-Modeling Perspective." *Astrobiology* 23, no. 5 (May): 496–512. https://doi.org/10.1089/ast.2022.0122.

———. 2024a. "Inferring Social Networks from Unstructured Text Data: A Proof of Concept Detection of Hidden Communities of Interest." *Data & Policy* 6:e5. https://doi.org/10.1017/dap.2023.48.

———. 2024b. "Visualizing Hidden Communities of Interest: A Case-Study Analysis of Topic-Based Social Networks in Astrobiology." *Scientometrics* (May). https://doi.org/10.1007/s11192-024-05047-7.

Malaterre, Christophe, Francis Lareau, Davide Pulizzotto, and Jonathan St-Onge. 2021. "Eight Journals over Eight Decades: A Computational Topic-Modeling Approach to Contemporary Philosophy of Science." *Synthese* 199, nos. 1-2 (December): 2883–2923. https://doi.org/10.1007/s11229-020-02915-6.

Malaterre, Christophe, and Martin Léonard. 2023. "Charting the Territories of Epistemic Concepts in the Practice of Science: A Text-Mining Approach." *The British Journal for the Philosophy of Science* (April): 725656. https://doi.org/10.1086/725656.

———. 2024. "Epistemic Markers in the Scientific Discourse." *Philosophy of Science* 91, no. 1 (January): 151–174. https://doi.org/10.1017/psa.2023.97.

Malaterre, Christophe, Davide Pulizzotto, and Francis Lareau. 2020. "Revisiting Three Decades of Biology and Philosophy: A Computational Topic-Modeling Perspective." *Biology & Philosophy* 35, no. 1 (February): 5. https://doi.org/10.1007/s10539-019-9729-4.

Malinowski, Bronislaw. 2002. *Argonauts of the Western Pacific: An Account of Native Enterprise and Adventure in the Archipelagoes of Melanesian New Guinea.* Hoboken: Taylor and Francis.

Mamak, Wojciech, and Marcin Milkowski. 2025. "Meaning from Matching: How Representational Mechanisms Exploit Structural Similarity." *THEORIA. An International Journal for Theory, History and Foundations of Science* (May). https://doi.org/10.1387/theoria.25171.

Man, Kingson, and Antonio R. Damasio. 2019. "Homeostasis and Soft Robotics in the Design of Feeling Machines." *Nature Machine Intelligence* 1, no. 10 (October): 446–452. https://doi.org/10.1038/s42256-019-0103-7.

Mann, Stephen Francis. 2018. "Attribution of Information in Animal Interaction." *Biological Theory* 13, no. 3 (September): 164–179. https://doi.org/10.1007/s13752-018-0299-5.

———. 2020. "Consequences of a Functional Account of Information." *Review of Philosophy and Psychology* 11, no. 3 (September): 669–687. https://doi.org/10.1007/s13164-018-0413-4.

Mannella, Francesco, and Luca Tummolini. 2023. "Kick-Starting Concept Formation with Intrinsically Motivated Learning: The Grounding by Competence Acquisition Hypothesis." *Philosophical Transactions of the Royal Society B: Biological Sciences* 378, no. 1870 (February): 20210370. https://doi.org/10.1098/rstb.2021.0370.

Margulis, Lynn. 1991. "Symbiogenesis and Symbionticism." In *Symbiosis as a Source of Evolutionary Innovation: Speciation and Morphogenesis,* edited by Lynn Margulis and René Fester, 1–14. Cambridge, Mass: MIT Press.

Marr, David. [2010]1982. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.* Cambridge, Mass: MIT Press.

Massimi, Michela. 2022. *Perspectival Realism.* New York: Oxford University Press.

Massimi, Michela, and Casey D. McCoy. 2019. *Understanding Perspectivism.* New York: Routledge, June. https://doi.org/10.4324/9781315145198-9.

Maturana, Humberto R., and Francisco J. Varela. 1980. *Autopoiesis and Cognition: The Realization of the Living.* Vol. 42. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-009-8947-4.

Maynard Smith, John. 2000. "The Concept of Information in Biology." *Philosophy of Science* 67, no. 2 (June): 177–194. https://doi.org/10.1086/392768.

Maynard Smith, John, and David Harper. 2003. *Animal Signals.* New York: Oxford University Press.

Mayr, Ernst. 1961. "Cause and Effect in Biology." *Science* 134 (3489): 1501–1506.

———. 1992. "The Idea of Teleology." *Journal of the History of Ideas* 53 (1): 117–135. https://doi.org/10.2307/2709913.

McAlister, Simon, Colin Allen, Andrew Ravenscroft, Chris Reed, David Bourget, John Lawrence, Katy Börner, and Robert Light. 2014. *From Big Data to Argument Analysis and Automated Extraction: A Selective Study of Argument in the Philosophy of Animal Psychology from the Volumes of the Hathi Trust Collection.* Technical report. April. Accessed September 25, 2024.

McCaffrey, Joseph, and Jessey Wright. 2022. "Neuroscience and Cognitive Ontology: A Case for Pluralism." In *Neuroscience and Philosophy,* edited by Felipe De Brigard and Walter Sinnott-Armstrong, 427–466. The MIT Press, February. https://doi.org/10.7551/mitpress/12611.003.0019.

McCulloch, Warren S., and Walter Pitts. 1943. "A Logical Calculus of the Ideas Immanent in Nervous Activity." *The Bulletin of Mathematical Biophysics* 5, no. 4 (December): 115–133. https://doi.org/10.1007/BF02478259.

McGann, Marek. 2024. "Facing Life: The Messy Bodies of Enactive Cognitive Science." *Phenomenology and the Cognitive Sciences* (January). https://doi.org/10.1007/s11097-024-09958-x.

McGivern, Patrick. 2020. "Active Materials: Minimal Models of Cognition?" *Adaptive Behavior* 28, no. 6 (December): 441–451. https://doi.org/10.1177/1059712319891742.

McInnes, Leland, John Healy, and James Melville. 2020. *UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction,* arXiv:1802.03426, September. Accessed June 20, 2025. https://doi.org/10.48550/arXiv.1802.03426.

McMillen, Patrick, and Michael Levin. 2024. "Collective Intelligence: A Unifying Concept for Integrating Biology across Scales and Substrates." *Communications Biology* 7, no. 1 (March): 378. https://doi.org/10.1038/s42003-024-06037-4.

McNally, Louise. 2013. "Semantics and Pragmatics." *WIREs Cognitive Science* 4, no. 3 (May): 285–297. https://doi.org/10.1002/wcs.1227.

Meacham, Frazer, Aaron Perlmutter, and Carl T. Bergstrom. 2013. "Honest Signalling with Costly Gambles." *Journal of The Royal Society Interface* 10, no. 87 (October): 20130469. https://doi.org/10.1098/rsif.2013.0469.

Medaglia, John D., Deepa M. Ramanathan, Umesh M. Venkatesan, and Frank G. Hillary. 2011. "The Challenge of Non-Ergodicity in Network Neuroscience." *Network: Computation in Neural Systems* 22, nos. 1-4 (March): 148–153. https://doi.org/10.3109/09638237.2011.639604.

Medawar, Peter. 1996. "Is the Scientific Paper a Fraud?" In *The Strange Case of the Spotted Mice and Other Classic Essays on Science,* 33–39. Oxford University Press, April. Accessed November 12, 2024. https://doi.org/10.1093/oso/9780192861931.003.0003.

Melander, Peter. 1993. "How Not to Explain the Errors of the Immune System." *Philosophy of Science* 60 (2): 223–241.

Merleau-Ponty, Maurice. 2005. *Phenomenology of Perception.* Translated by Colin Smith. London: Routledge.

Meynell, Letitia. 2013. "Parsing Pictures: On Analyzing the Content of Images in Science." *The Knowledge Engineering Review* 28, no. 3 (September): 327–345. https://doi.org/10.1017/S0269888913000271.

Michel, Jean-Baptiste, Yuan Kui Shen, Aviva Presser Aiden, Adrian Veres, Matthew K. Gray, The Google Books Team, Joseph P. Pickett, et al. 2011. "Quantitative Analysis of Culture Using Millions of Digitized Books." *Science* 331, no. 6014 (January): 176–182. https://doi.org/10.1126/science.1199644.

Mickus, Timothee, Denis Paperno, Mathieu Constant, and Kees van Deemter. 2020. "What Do You Mean, BERT? Assessing BERT as a Distributional Semantics Model." In *Proceedings of the Society for Computation in Linguistics 2020,* 279–290. New York, New York: Association for Computational Linguistics. https://doi.org/10.7275/T778-JA71.

Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. *Efficient Estimation of Word Representations in Vector Space,* arXiv:1301.3781, September. Accessed June 3, 2022.

Miłkowski, Marcin. 2013. *Explaining the Computational Mind.* Cambridge, Massachusetts: The MIT Press.

———. 2020. "Thinking about Semantic Information." *AVANT. The Journal of the Philosophical-Interdisciplinary Vanguard* 11 (2). https://doi.org/10.26913/avant.2020.02.08.

———. 2022. "Cognitive Artifacts and Their Virtues in Scientific Practice." *Studies in Logic, Grammar and Rhetoric* 67, no. 3 (December): 219–246. https://doi.org/10.2478/slgr-2022-0012.

———. 2023a. "Cognitive Metascience: A New Approach to the Study of Theories." *Przegląd Psychologiczny* 66, no. 1 (October): 185–207. https://doi.org/10.31648/przegldpsychologiczny.9682.

———. 2023b. "Correspondence Theory of Semantic Information." *The British Journal for the Philosophy of Science* 74, no. 2 (June): 485–510. https://doi.org/10.1086/714804.

———. 2025. "Interactivism Mechanized: Bridging the Gap between Cognition, Correspondence, and Computation." *Phenomenology and the Cognitive Sciences* (June). https://doi.org/10.1007/s11097-025-10082-7.

Miłkowski, Marcin, Robert Clowes, Zuzanna Rucińska, Aleksandra Przegalińska, Tadeusz Zawidzki, Joel Krueger, Adam Gies, et al. 2018. "From Wide Cognition to Mechanisms: A Silent Revolution." *Frontiers in Psychology* 9:2393. https://doi.org/10.3389/fpsyg.2018.02393.

Miłkowski, Marcin, Witold M. Hensel, and Mateusz Hohol. 2018. "Replicability or Reproducibility? On the Replication Crisis in Computational Neuroscience and Sharing Only Relevant Detail." *Journal of Computational Neuroscience* 45, no. 3 (December): 163–172. https://doi.org/10.1007/s10827-018-0702-z.

Miłkowski, Marcin, and Mateusz Hohol. 2021. "Explanations in Cognitive Science: Unification versus Pluralism." *Synthese* 199, no. S1 (November): 1–17. https://doi.org/10.1007/s11229-020-02777-y.

Miłkowski, Marcin, and Ksawery Jasieński. 2022. *eLife Open Peer Review Corpus.* Accessed October 3, 2023. https://doi.org/10.18150/FKPEQN.

Miłkowski, Marcin, and Przemysław Nowakowski. 2025. "Tracking Norms: A Plea for a Normative Digital Philosophy of Science." *Synthese* 206, no. 3 (August): 120. https://doi.org/10.1007/s11229-025-05200-6.

Miller, George A. 2003. "The Cognitive Revolution: A Historical Perspective." *Trends in Cognitive Sciences* 7, no. 3 (March): 141–144. https://doi.org/10.1016/s1364-6613(03)00029-9.

———. 1956. "The Magical Number Seven, plus or Minus Two: Some Limits on Our Capacity for Processing Information." *Psychological Review* 63, no. 2 (March): 81–97. https://doi.org/10.1037/h0043158.

Millikan, Ruth Garrett. 1984. *Language, Thought, and Other Biological Categories: New Foundations for Realism.* Cambridge, Mass.: MIT Press.

———. 1989a. "An Ambiguity in the Notion "Function"." *Biology & Philosophy* 4, no. 2 (April): 172–176. https://doi.org/10.1007/BF00127747.

———. 1989b. "Biosemantics." *The Journal of Philosophy* 86, no. 6 (June): 281. https://doi.org/10.2307/2027123.

———. 1989c. "In Defense of Proper Functions." *Philosophy of Science* 56, no. 2 (June): 288–302. https://doi.org/10.1086/289488.

———. 1995. "Pushmi-Pullyu Representations." *Philosophical Perspectives* 9:185. https://doi.org/10.2307/2214217.

———. 2004. *Varieties of Meaning.* 2002. Cambridge, Mass.: MIT Press.

———. 2012. "Natural Signs." In *How the World Computes,* edited by David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell, Moni Naor, et al., 7318:496–506. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-30870-3_50.

Mitchell, Melanie. 2019. *Artificial Intelligence: A Guide for Thinking Humans.* New York: Farrar, Straus and Giroux.

Mitchell, Sandra D. 2002. "Integrative Pluralism." *Biology and Philosophy* 17, no. 1 (January): 55–70. https://doi.org/10.1023/A:1012990030867.

———. 2003. *Biological Complexity and Integrative Pluralism.* Cambridge, UK; New York, N.Y.: Cambridge University Press.

———. 2009. *Unsimple Truths: Science, Complexity, and Policy.* Chicago; London: The University of Chicago Press Chicago and London.

Mizrahi, Moti. 2020. "The Case Study Method in Philosophy of Science: An Empirical Study." *Perspectives on Science* 28, no. 1 (February): 63–88. https://doi.org/10.1162/posc_a_00333.

Moksness, Lars, and Svein Ottar Olsen. 2017. "Understanding Researchers' Intention to Publish in Open Access Journals." *Journal of Documentation* 73, no. 6 (October): 1149–1166. https://doi.org/10.1108/JD-02-2017-0019.

Monod, Jacques. 1971. *Chance and necessity: an essay on the natural philosophy of modern biology.* New York: Knopf.

Morange, Michel. 2020. *The Black Box of Biology: A History of the Molecular Revolution.* Translated by Matthew Cobb. Cambridge: Harvard University Press.

Mordatch, Igor, and Pieter Abbeel. 2018. "Emergence of Grounded Compositional Language in Multi-Agent Populations." In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI Symposium on Educational Advances in Artificial Intelligence.* AAAI'18/IAAI'18/EAAI'18. New Orleans, Louisiana, USA: AAAI Press.

Moreno, Alvaro, and Matteo Mossio. 2015. *Biological Autonomy.* Vol. 12. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-017-9837-2.

Moretti, Franco. 2013. *Distant Reading.* London ; New York: Verso.

Morgan, C. Lloyd. 1894. *An Introduction to Comparative Psychology.* London: Walter Scott Publishing Co. https://doi.org/10.1037/11344-000.

Moser, Edvard I., Emilio Kropff, and May-Britt Moser. 2008. "Place Cells, Grid Cells, and the Brain's Spatial Representation System." *Annual Review of Neuroscience* 31, no. 1 (July): 69–89. https://doi.org/10.1146/annurev.neuro.31.061307.090723.

Moss, Lenny. 2009. "Detachment, Genomics and the Nature of Being Human." In *New Visions of Nature: Complexity and Authenticity,* edited by Martin A. M. Drenthen, F.W. Jozef Keulartz, and James Proctor, 103–115. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-90-481-2611-8_8.

Mossio, Matteo, Leonardo Bich, and Alvaro Moreno. 2013. "Emergence, Closure and Inter-level Causation in Biological Systems." *Erkenntnis* 78, no. S2 (December): 153–178. https://doi.org/10.1007/s10670-013-9507-7.

Mossio, Matteo, Cristian Saborido, and Alvaro Moreno. 2009. "An Organizational Account of Biological Functions." *The British Journal for the Philosophy of Science* 60, no. 4 (December): 813–841. https://doi.org/10.1093/bjps/axp036.

Mountcastle, V. 1997. "The Columnar Organization of the Neocortex." *Brain* 120, no. 4 (April): 701–722. https://doi.org/10.1093/brain/120.4.701.

Muller, Scott J. 2007. *Asymmetry: The Foundation of Information.* Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-69884-5.

Muñoz, Roberto N., Angus Leung, Aidan Zecevik, Felix A. Pollock, Dror Cohen, Bruno Van Swinderen, Naotsugu Tsuchiya, and Kavan Modi. 2020. "General Anesthesia Reduces Complexity and Temporal Asymmetry of the Informational Structures Derived from Neural Recordings in *Drosophila.*" *Physical Review Research* 2, no. 2 (May): 023219. https://doi.org/10.1103/PhysRevResearch.2.023219.

Murdock, Jaimie, Colin Allen, Katy Börner, Robert Light, Simon McAlister, Andrew Ravenscroft, Robert Rose, et al. 2017. "Multi-Level Computational Methods for Interdisciplinary Research in the HathiTrust Digital Library." *PLOS ONE* 12, no. 9 (September): e0184188. https://doi.org/10.1371/journal.pone.0184188.

Murdock, Jaimie, Colin Allen, and Simon DeDeo. 2017. "Exploration and Exploitation of Victorian Science in Darwin's Reading Notebooks." *Cognition* 159 (February): 117–126. https://doi.org/10.1016/j.cognition.2016.11.012.

Murphy, Gregory L. 2004. *The Big Book of Concepts.* Cambridge, Mass.: MIT Press.

Muthukrishna, Michael, and Joseph Henrich. 2019. "A Problem in Theory." *Nature Human Behaviour* 3, no. 3 (February): 221–229. https://doi.org/10.1038/s41562-018-0522-1.

Nagel, Ernest. 1961. *The Structure of Science.* Indianapolis, IN: Hackett.

Nanay, Bence. 2010. "A Modal Theory of Function." *The Journal of Philosophy* 107 (8): 412–431.

———. 2014. "Teleosemantics without Etiology." *Philosophy of Science* 81, no. 5 (December): 798–810. https://doi.org/10.1086/677684.

Naselaris, Thomas, Kendrick N. Kay, Shinji Nishimoto, and Jack L. Gallant. 2011. "Encoding and Decoding in fMRI." *NeuroImage* 56, no. 2 (May): 400–410. https://doi.org/10.1016/j.neuroimage.2010.07.073.

Nasuto, Slawomir J., Kerstin Dautenhahn, and Mark Bishop. 1999. "Communication as an Emergent Metaphor for Neuronal Operation." In *Computation for Metaphors, Analogy, and Agents,* edited by G. Goos, J. Hartmanis, J. Van Leeuwen, and Chrystopher L. Nehaniv, 1562:365–379. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-48834-0_19.

Nazir, Shahzad, Muhammad Asif, Shahbaz Ahmad, Faisal Bukhari, Muhammad Tanvir Afzal, and Hanan Aljuaid. 2020. "Important Citation Identification by Exploiting Content and Section-Wise in-Text Citation Count." Edited by Wolfgang Glanzel. *PLOS ONE* 15, no. 3 (March): e0228885. https://doi.org/10.1371/journal.pone.0228885.

Neander, Karen. 1991. "The Teleological Notion of 'Function'." *Australasian Journal of Philosophy* 69, no. 4 (December): 454–468. https://doi.org/10.1080/00048409112344881.

———. 2012. "Toward an Informational Teleosemantics." In *Millikan and Her Critics,* 1st ed., edited by Dan Ryder, Justine Kingsbury, and Kenneth Williford, 21–40. Wiley, November. https://doi.org/10.1002/9781118328118.ch1.

———. 2017. *A Mark of the Mental: In Defense of Informational Teleosemantics.* Cambridge, Massachusetts London, England: The MIT Press.

Nelson, Nicole C. 2018. *Model Behavior: Animal Experiments, Complexity, and the Genetics of Psychiatric Disorders.* Chicago: The University of Chicago Press.

Nemati, Nedah. 2024. "Rethinking Neuroscientific Methodology: Lived Experience in Behavioral Studies." *Biological Theory* 19, no. 3 (September): 184–197. https://doi.org/10.1007/s13752-024-00460-w.

Nersessian, Nancy J. 2010. *Creating Scientific Concepts.* Cambridge, Mass.: MIT Press.

Neurath, Otto. 1937. "Unified Science and Its Encyclopaedia." *Philosophy of Science* 4, no. 2 (April): 265–277. https://doi.org/10.1086/286457.

Newell, Allen, J. C. Shaw, and Herbert A. Simon. 1958. "Elements of a Theory of Human Problem Solving." *Psychological Review* 65 (3): 151–166. https://doi.org/10.1037/h0048495.

Newell, Allen, and Herbert A. Simon. 1972. *Human Problem Solving.* Englewood Cliffs, NJ: Prentice-Hall, Inc.

Nomikou, Iris, Giuseppe Leonardi, Alicja Radkowska, Joanna Rączaszek-Leonardi, and Katharina J. Rohlfing. 2017. "Taking Up an Active Role: Emerging Participation in Early Mother–Infant Interaction during Peekaboo Routines." *Frontiers in Psychology* 8 (October): 1656. https://doi.org/10.3389/fpsyg.2017.01656.

Nomikou, Iris, Giuseppe Leonardi, Karharina J. Rohlfing, and Joanna Rączaszek-Leonardi. 2016. "Constructing Interaction: The Development of Gaze Dynamics." *Infant and Child Development* 25, no. 3 (May): 277–295. https://doi.org/10.1002/icd.1975.

Norman, Donald A. 1991. "Cognitive Artifacts." In *Designing Interaction: Psychology at the Human-Computer Interface.* Edited by John M. Carroll, 17–38. Cambridge Series on Human-Computer Interaction, No. 4. New York, NY, US: Cambridge University Press.

Nowak, Leszek. 1972. "Laws of Science, Theories, Measurement: (Comments on Ernest Nagel's *the Structure of Science*)." *Philosophy of Science* 39, no. 4 (December): 533–548. https://doi.org/10.1086/288479.

———. 1980. *The Structure of Idealization: Towards a Systematic Interpretation of the Marxian Idea of Science.* Dordrecht ; Boston: D. Reidel.

Nowak, Martin A., Joshua B. Plotkin, and Vincent A. A. Jansen. 2000. "The Evolution of Syntactic Communication." *Nature* 404, no. 6777 (March): 495–498. https://doi.org/10.1038/35006635.

O'Connor, Cailin. 2020. *Games in the Philosophy of Biology.* Cambridge University Press, January. https://doi.org/10.1017/9781108616737.

O'Keefe, J., N. Burgess, J. G. Donnett, K. J. Jeffery, and E. A. Maguire. 1998. "Place Cells, Navigational Accuracy, and the Human Hippocampus." *Philosophical Transactions of the Royal Society B: Biological Sciences* 353, no. 1373 (August): 1333–1340. https://doi.org/10.1098/rstb.1998.0287.

Oberauer, Klaus, and Stephan Lewandowsky. 2019. "Addressing the Theory Crisis in Psychology." *Psychonomic Bulletin & Review* 26, no. 5 (October): 1596–1618. https://doi.org/10.3758/s13423-019-01645-2.

Okasha, Samir. 2008. *Evolution and the Levels of Selection.* Oxford : Oxford ; New York: Clarendon Press ; Oxford University Press.

Open Science Collaboration. 2015. "Estimating the Reproducibility of Psychological Science." *Science* 349, no. 6251 (August): aac4716. https://doi.org/10.1126/science.aac4716.

OpenAI. 2023. *GPT-4 Technical Report,* arXiv:2303.08774, March. Accessed May 25, 2023.

Oppenheim, Paul, and Hilary Putnam. 1958. "The Unity of Science as a Working Hypothesis." *Minnesota Studies in the Philosophy of Science* 2:3–36.

Oreskes, Naomi. 2021. *Why Trust Science?* Princeton, New Jersey: Princeton University Press.

Oreskes, Naomi, and Erik M. Conway. 2010. *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming.* New York: Bloomsbury Press.

Osbeck, Lisa M., and Nancy J. Nersessian. 2014. "Situating Distributed Cognition." *Philosophical Psychology* 27, no. 1 (January): 82–97. https://doi.org/10.1080/09515089.2013.829384.

Overton, James A. 2013. ""Explain" in Scientific Discourse." *Synthese* 190, no. 8 (May): 1383–1405. https://doi.org/10.1007/s11229-012-0109-8.

Oviedo, Néstor J., Junji Morokuma, Peter Walentek, Ido P. Kema, Man Bock Gu, Joo-Myung Ahn, Jung Shan Hwang, Takashi Gojobori, and Michael Levin. 2010. "Long-Range Neural and Gap Junction Protein-Mediated Cues Control Polarity during Planarian Regeneration." *Developmental Biology* 339, no. 1 (March): 188–199. https://doi.org/10.1016/j.ydbio.2009.12.012.

Owren, Michael J., Drew Rendall, and Michael J. Ryan. 2010. "Redefining Animal Signaling: Influence versus Information in Communication." *Biology & Philosophy* 25, no. 5 (November): 755–780. https://doi.org/10.1007/s10539-010-9224-4.

Oyama, Susan. 2000. *The Ontogeny of Information: Developmental Systems and Evolution.* Durham, NC: Duke Univ. Press.

Pankevich, Diana E., Bruce M. Altevogt, Miriam Davis, and Institute of Medicine (U.S.), eds. 2011. *Glutamate-Related Biomarkers in Drug Development for Disorders of the Nervous System: Workshop Summary.* Washington, D.C: National Academies Press.

Pattee, Howard H. (1969) 2012. "How Does a Molecule Become a Message?" In *Laws, Language and Life,* 7:55–67. Dordrecht: Springer Netherlands. Accessed June 9, 2020. https://doi.org/10.1007/978-94-007-5161-3_3.

———. (1972) 2012. "Laws and Constraints, Symbols and Languages." In *Laws, Language and Life,* 7:81–89. Dordrecht: Springer Netherlands. Accessed April 22, 2025. https://doi.org/10.1007/978-94-007-5161-3_5.

———, ed. 1973a. *Hierarchy Theory: The Challenge of Complex Systems.* New York: Braziller.

———. (1973b) 2012. "Physical Problems of Decision-Making Constraints." In *Laws, Language and Life,* 7:69–79. Dordrecht: Springer Netherlands. Accessed April 22, 2025. https://doi.org/10.1007/978-94-007-5161-3_4.

———. (1973c) 2012. "The Physical Basis and Origin of Hierarchical Control." In *Laws, Language and Life,* 7:91–110. Dordrecht: Springer Netherlands. Accessed April 22, 2025. https://doi.org/10.1007/978-94-007-5161-3_6.

———. (1982) 2012. "Cell Psychology: An Evolutionary Approach to the Symbol-Matter Problem." In *Laws, Language and Life,* 7:165–179. Dordrecht: Springer Netherlands. Accessed June 13, 2021. https://doi.org/10.1007/978-94-007-5161-3_11.

———. 1997. "The Physics of Symbols and the Evolution of Semiotic Controls." In *Proceedings of the Workshop on Control Mechanisms for Complex Systems: Issues of Measurement and Semiotic Analysis,* edited by M. Coombs and M. Sulcoski. Las Cruces: University of New Mexico Press.

———. 2012. "Introduction—What These Papers Are About." In *Laws, Language and Life,* 7:3–30. Dordrecht: Springer Netherlands. Accessed April 22, 2025. https://doi.org/10.1007/978-94-007-5161-3_1.

———. 2017. "The Problem of Biological Hierarchy." In *Organization Stability & Process,* 3rd ed., 117–136. Routledge, July. Accessed April 22, 2025. https://doi.org/10.4324/9781315125893-10.

———. 2021. "Symbol Grounding Precedes Interpretation." *Biosemiotics* 14, no. 3 (December): 561–568. https://doi.org/10.1007/s12304-021-09458-4.

———. 2023. "Symbol Grounding by Folding: The Primary Biosemiosis." In *Open Semiotics,* edited by Amir Biglari. Paris: L'Harmattan.

Pearl, Judea. 2022. *Causality: Models, Reasoning, and Inference.* Cambridge; New York, NY; Port Melbourne; New Delhi; Singapore: Cambridge University Press.

Peirce, Charles S. 1985. *Collected Papers of Charles Sanders Peirce: Principles of Philosophy and Elements of Logic.* Edited by Charles Hartshorne and Paul Weiss. Cambridge, Mass.: Belknap Press of Harvard Univ. Press.

Pence, Charles H. 2022a. "How Not to Fight About Theory: The Debate Between Biometry and Mendelism in "Nature", 1890–1915." In *The Dynamics of Science: Computational Frontiers in History and Philosophy of Science,* edited by Grant Ramsey and Andreas De Block. Pittsburgh, Pa: University of Pittsburgh press.

———. 2022b. "Testing and Discovery: Responding to Challenges to Digital Philosophy of Science." *Metaphilosophy* 53, nos. 2-3 (April): 238–253. https://doi.org/10.1111/meta.12549.

———. 2025. "Textual Analysis and Conceptual Cartography." In *Methods in Philosophy of Science: A User's Guide,* edited by Adrian Currie and Sophie Veigl. MIT Press.

Perini, Laura. 2005. "The Truth in Pictures." *Philosophy of Science* 72, no. 1 (January): 262–285. https://doi.org/10.1086/426852.

Perkel, Donald H., and Theodore Holmes Bullock. 1968. *Neural Coding: A Report Based on an NRP Work Session.* Technical report 6 (3). Brookline, MA: Neuroscience Research Program, December.

Pero, Francesca. 2024. "Target Systems." In *The Routledge Handbook of Philosophy of Scientific Modeling,* 1st ed., edited by Tarja Knuuttila, Natalia Carrillo, and Rami Koskinen, 126–137. London: Routledge, August. https://doi.org/10.4324/9781003205647.

Pezzulo, Giovanni, Lawrence W. Barsalou, Angelo Cangelosi, Martin H. Fischer, Ken McRae, and Michael J. Spivey. 2013. "Computational Grounded Cognition: A New Alliance between Grounded Cognition and Computational Modeling." *Frontiers in Psychology* 3. https://doi.org/10.3389/fpsyg.2012.00612.

Piccinini, Gualtiero. 2020. "Nonnatural Mental Representation." In *What Are Mental Representations?,* 1st ed., edited by Joulia Smortchkova, Krzysztof Dołęga, and Tobias Schlicht, 254–286. New York: Oxford University Press, December. https://doi.org/10.1093/oso/9780190686673.003.0010.

———. 2021. *Neurocognitive Mechanisms: Explaining Biological Cognition.* Oxford: Oxford University Press.

Pichler, Axel, and Nils Reiter. 2022. "From Concepts to Texts and Back: Operationalization as a Core Activity of Digital Humanities." *Journal of Cultural Analytics* 7, no. 4 (December). https://doi.org/10.22148/001c.57195.

Pickering, Andrew. 2010. *The Cybernetic Brain: Sketches of Another Future.* Chicago ; London: University of Chicago Press.

Pigliucci, Massimo, and Gerd Müller. 2010a. "Elements of an Extended Evolutionary Synthesis." In *Evolution, the Extended Synthesis,* edited by Massimo Pigliucci and Gerd Müller, 3–17. Cambridge, Mass: MIT Press.

———, eds. 2010b. *Evolution, the Extended Synthesis.* Cambridge, Mass: MIT Press.

Pinotsis, Dimitris A, and Earl K Miller. 2023. "In Vivo Ephaptic Coupling Allows Memory Network Formation." *Cerebral Cortex* 33, no. 17 (August): 9877–9895. https://doi.org/10.1093/cercor/bhad251.

Pisula, Wojciech. 2016. "Levels of Consciousness." *Open Journal of Philosophy* 06 (01): 51–58. https://doi.org/10.4236/ojpp.2016.61006.

Pitt, Joseph C. 2001. "The Dilemma of Case Studies: Toward a Heraclitian Philosophy of Science." *Perspectives on Science* 9, no. 4 (December): 373–382. https://doi.org/10.1162/106361401760375785.

Piwowar, Heather, Jason Priem, Vincent Larivière, Juan Pablo Alperin, Lisa Matthias, Bree Norlander, Ashley Farley, Jevin West, and Stefanie Haustein. 2018. "The State of OA: A Large-Scale Analysis of the Prevalence and Impact of Open Access Articles." *PeerJ* 6 (February): e4375. https://doi.org/10.7717/peerj.4375.

Planer, Ronald J. 2014. "Replacement of the "Genetic Program" Program." *Biology & Philosophy* 29, no. 1 (January): 33–53. https://doi.org/10.1007/s10539-013-9388-9.

Pocheville, Arnaud. 2018. "Biological Information as Choice and Construction." *Philosophy of Science* 85, no. 5 (December): 1012–1025. https://doi.org/10.1086/699832.

Polger, Thomas W. 2004. *Natural Minds.* Cambridge, Mass.: MIT Press.

Polger, Thomas W., and Lawrence A. Shapiro. 2016. *The Multiple Realization Book.* Oxford: Oxford University Press.

Potochnik, Angela. 2011. "A Neurathian Conception of the Unity of Science." *Erkenntnis* 74, no. 3 (May): 305–319. https://doi.org/10.1007/s10670-010-9228-0.

———. 2017. *Idealization and the Aims of Science.* Chicago: The University of Chicago Press.

Potter, Henry D., and Kevin J. Mitchell. 2022. "Naturalising Agent Causation." *Entropy* 24, no. 4 (March): 472. https://doi.org/10.3390/e24040472.

Pradeu, Thomas, Maël Lemoine, Mahdi Khelfaoui, and Yves Gingras. 2024. "Philosophy in Science: Can Philosophers of Science Permeate through Science and Produce Scientific Knowledge?" *The British Journal for the Philosophy of Science* 75, no. 2 (June): 375–416. https://doi.org/10.1086/715518.

Price, Tabitha, Philip Wadewitz, Dorothy Cheney, Robert Seyfarth, Kurt Hammerschmidt, and Julia Fischer. 2015. "Vervets Revisited: A Quantitative Analysis of Alarm Call Structure and Context Specificity." *Scientific Reports* 5, no. 1 (August): 13220. https://doi.org/10.1038/srep13220.

Prindle, Arthur, Jintao Liu, Munehiro Asally, San Ly, Jordi Garcia-Ojalvo, and Gürol M. Süel. 2015. "Ion Channels Enable Electrical Communication in Bacterial Communities." *Nature* 527, no. 7576 (November): 59–63. https://doi.org/10.1038/nature15709.

Rączaszek-Leonardi, Joanna. 2012. "Language as a System of Replicable Constraints." In *Laws, Language and Life,* 7:295–333. Dordrecht: Springer Netherlands. Accessed June 6, 2022. https://doi.org/10.1007/978-94-007-5161-3_19.

———. 2016. "How Does a Word Become a Message? An Illustration on a Developmental Time-Scale." *New Ideas in Psychology* 42 (August): 46–55. https://doi.org/10.1016/j.newideapsych.2015.08.001.

———. 2023. "What Dynamic Approaches Have Taught Us About Cognition and What They Have Not: On Values in Motion and the Importance of Replicable Forms." *Topics in Cognitive Science* (November): tops.12709. https://doi.org/10.1111/tops.12709.

Rączaszek-Leonardi, Joanna, and Terrence W. Deacon. 2018. “Ungrounding Symbols in Language Development: Implications for Modeling Emergent Symbolic Communication in Artificial Systems.” In *2018 Joint IEEE 8th International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob),* 232–237. Tokyo, Japan: IEEE, September. https://doi.org/10.1109/DEVLRN.2018.8761016.

Rączaszek-Leonardi, Joanna, Krzysztof Główka, Iris Nomikou, and Nicole Rossmanith. 2022. “Time-to-Smile, Time-to-Speak, Time-to-Resolve: Timescales for Shaping Engagement in Language.” *Language Sciences* 93 (September): 101495. https://doi.org/10.1016/j.langsci.2022.101495.

Rączaszek-Leonardi, Joanna, Iris Nomikou, and Katharina J. Rohlfing. 2013. “Young Children’s Dialogical Actions: The Beginnings of Purposeful Intersubjectivity.” *IEEE Transactions on Autonomous Mental Development* 5, no. 3 (September): 210–221. https://doi.org/10.1109/TAMD.2013.2273258.

Rączaszek-Leonardi, Joanna, Iris Nomikou, Katharina J. Rohlfing, and Terrence W. Deacon. 2018. “Language Development From an Ecological Perspective: Ecologically Valid Ways to Abstract Symbols.” *Ecological Psychology* 30, no. 1 (January): 39–73. https://doi.org/10.1080/10407413.2017.1410387.

Rączaszek-Leonardi, Joanna, and Julian Zubek. 2023. “Is Love an Abstract Concept? A View of Concepts from an Interaction-Based Perspective.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 378, no. 1870 (February): 20210356. https://doi.org/10.1098/rstb.2021.0356.

Raja, Vicente. 2018. “A Theory of Resonance: Towards an Ecological Cognitive Architecture.” *Minds and Machines* 28, no. 1 (March): 29–51. https://doi.org/10.1007/s11023-017-9431-8.

———. 2021. “Resonance and Radical Embodiment.” *Synthese* 199, no. 1 (November): 113–141. https://doi.org/10.1007/s11229-020-02610-6.

———. 2024. “The Motifs of Radical Embodied Neuroscience.” *European Journal of Neuroscience* 60, no. 5 (September): 4738–4755. https://doi.org/10.1111/ejn.16434.

Raman, Anantanarayanan. 2011. “Morphogenesis of Insect-Induced Plant Galls: Facts and Questions.” *Flora - Morphology, Distribution, Functional Ecology of Plants* 206, no. 6 (June): 517–533. https://doi.org/10.1016/j.flora.2010.08.004.

Récanati, François. 2013. *Mental Files.* Oxford: Oxford University Press.

Reichenbach, Hans. 1938. *Experience and Prediction.* Chicago: The University of Chicago Press.

———. 1958. *The Philosophy of Space and Time.* Translated by Maria Reichenbach and John Freund. New York: Dover Publications.

Reimers, Nils, and Iryna Gurevych. 2019. “Sentence-BERT: Sentence Embeddings Using Siamese BERT-networks.” In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing.* Association for Computational Linguistics, November.

Rendall, Drew, Michael J. Owren, and Michael J. Ryan. 2009. “What Do Animal Signals Mean?” *Animal Behaviour* 78, no. 2 (August): 233–240. https://doi.org/10.1016/j.anbehav.2009.06.007.

Reutlinger, Alexander. 2013. *A Theory of Causation in the Social and Biological Sciences.* London: Palgrave Macmillan UK. https://doi.org/10.1057/9781137281043.

Ritchie, J. Brendan, David Michael Kaplan, and Colin Klein. 2019. “Decoding the Brain: Neural Representation and the Limits of Multivariate Pattern Analysis in Cognitive Neuroscience.” *The British Journal for the Philosophy of Science* 70, no. 2 (June): 581–607. https://doi.org/10.1093/bjps/axx023.

Roberts, Gareth. 2017. “The Linguist’s *Drosophila*: Experiments in Language Change.” *Linguistics Vanguard* 3, no. 1 (August). https://doi.org/10.1515/lingvan-2016-0086.

Roberts, Gareth, and Bruno Galantucci. 2016. “Investigating Meaning in Experimental Semiotics.” *Psychology of Language and Communication* 20, no. 2 (December): 130–153. https://doi.org/10.1515/plc-2016-0008.

Robinson, Amanda K., Genevieve L. Quek, and Thomas A. Carlson. 2023. “Visual Representations: Insights from Neural Decoding.” *Annual Review of Vision Science* 9, no. 1 (September): 313–335. https://doi.org/10.1146/annurev-vision-100120-025301.

Robinson, Michael B., Meredith L. Lee, and Sabrina DaSilva. 2020. “Glutamate Transporters and Mitochondria: Signaling, Co-compartmentalization, Functional Coupling, and Future Directions.” *Neurochemical Research* 45, no. 3 (March): 526–540. https://doi.org/10.1007/s11064-020-02974-8.

Röder, Michael, Andreas Both, and Alexander Hinneburg. 2015. “Exploring the Space of Topic Coherence Measures.” In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining,* 399–408. Shanghai China: ACM, February. https://doi.org/10.1145/2684822.2685324.

Rorot, Wiktor. under review. *Looking for Keys in Model Organisms: Interpreting and Keying-up in Basal Cognition,* 27054.

Rorot, Wiktor, and Marcin Miłkowski. 2024. “Empirical Evidence in Conceptual Engineering, or the Defense of ‘Predictive Understanding’.” In *Conceptual Engineering,* edited by Piotr Stalmaszczyk, 255–280. Brill | mentis, September. https://doi.org/10.30965/9783969753026_013.

Rorot, Wiktor, and Joanna Rączaszek-Leonardi. 2023. “Understanding ”Compositionality” in Research on Language Emergence.” *Proceedings of the Annual Meeting of the Cognitive Science Society* 45.

Rorot, Wiktor, Katarzyna Skowrońska, Ewa Nagórska, Konrad Zielinski, Julian Zubek, and Joanna Rączaszek-Leonardi. 2022. *Structuring Unleashed Expression: Developmental Foundations of Human Communication.* Preprint. PsyArXiv, April. Accessed April 29, 2022. https://doi.org/10.31234/osf.io/tyva9.

Rosen, Robert. 1986. “SOME COMMENTS ON SYSTEMS AND SYSTEM THEORY.” *International Journal of General Systems* 13, no. 1 (November): 1–3. https://doi.org/10.1080/03081078608934949.

———. 1989. “The Roles of Necessity in Biology.” In *Newton to Aristotle,* edited by John Casti and Anders Karlqvist, 11–37. Boston, MA: Birkhäuser Boston. https://doi.org/10.1007/978-1-4684-0553-8_2.

———. 1991. *Life Itself: A Comprehensive Inquiry into the Nature, Origin, and Fabrication of Life.* New York: Columbia Univ. Press.

———. 2012. *Anticipatory Systems: Philosophical, Mathematical, and Methodological Foundations.* Vol. 1. New York, NY: Springer New York. https://doi.org/10.1007/978-1-4614-1269-4.

Rosenberg, Alexander, and Daniel W. McShea. 2012. *Philosophy of Biology: A Contemporary Introduction.* New York London: Routledge.

Ross, Lauren N. 2018. “Causal Selection and the Pathway Concept.” *Philosophy of Science* 85, no. 4 (October): 551–572. https://doi.org/10.1086/699022.

———. 2021a. “Causal Concepts in Biology: How Pathways Differ from Mechanisms and Why It Matters.” *The British Journal for the Philosophy of Science* 72, no. 1 (March): 131–158. https://doi.org/10.1093/bjps/axy078.

———. 2021b. “Causes with Material Continuity.” *Biology & Philosophy* 36, no. 6 (December): 52. https://doi.org/10.1007/s10539-021-09826-x.

———. 2021c. “Tracers in Neuroscience: Causation, Constraints, and Connectivity.” *Synthese* 199, nos. 1-2 (December): 4077–4095. https://doi.org/10.1007/s11229-020-02970-z.

———. 2025. *Explanation in Biology.* Cambridge University Press, January. https://doi.org/10.1017/9781009300940.

Rouse, Joseph. 1990. “The Narrative Reconstruction of Science.” *Inquiry* 33, no. 2 (January): 179–196. https://doi.org/10.1080/00201749008602217.

———. 2023. “The Radical Naturalism of Naturalistic Philosophy of Science.” *Topoi* 42, no. 3 (July): 719–732. https://doi.org/10.1007/s11245-023-09885-7.

Rowlands, Mark. 2010. *The New Science of the Mind: From Extended Mind to Embodied Phenomenology.* Cambridge, Mass: MIT Press.

Rychlý, Pavel. 2008. “A Lexicographer-Friendly Association Score.” In *Proceedings of Recent Advances in Slavonic Natural Language Processing,* edited by Petr Sojka and Aleš Horák, 6–9. Masaryk University, Brno.

Salmon, Wesley C. 1978. “Why Ask, ”Why?”? An Inquiry Concerning Scientific Explanation.” *Proceedings and Addresses of the American Philosophical Association* 51, no. 6 (August): 683. https://doi.org/10.2307/3129654.

———. 1984. *Scientific Explanation and the Causal Structure of the World.* Princeton, NJ: Princeton Univ. Pr.

———. 1994. “Causality without Counterfactuals.” *Philosophy of Science* 61 (2): 297–312.

Salton, Gerard, Edward A. Fox, and Harry Wu. 1983. “Extended Boolean Information Retrieval.” *Communications of the ACM* 26, no. 11 (November): 1022–1036. https://doi.org/10.1145/182.358466.

Sarkar, Sahotra. 1996. “Biological Information: A Skeptical Look at Some Central Dogmas of Molecular Biology.” *Boston Studies in the Philosophy of Science* 183:187–232.

———. 2000. “Information in Genetics and Developmental Biology: Comments on Maynard Smith.” *Philosophy of Science* 67 (2): 208–213. https://doi.org/10.1086/392771.

Sartori, Lorenzo. Forthcoming. “Why We Love Pictures (for the Wrong Reasons).” *Philosophy of Science.*

Saxena, Shreya, and John P Cunningham. 2019. “Towards the Neural Population Doctrine.” *Current Opinion in Neurobiology* 55 (April): 103–111. https://doi.org/10.1016/j.conb.2019.02.002.

Sayood, Khalid. 2018. “Information Theory and Cognition: A Review.” *Entropy (Basel, Switzerland)* 20, no. 9 (September): 706. https://doi.org/10.3390/e20090706.

Scarantino, Andrea. 2013. "Animal Communication as Information- Mediated Influence." In *Animal Communication Theory: Information and Influence,* edited by Ulrich E. Stegmann, 63–87. Cambridge ; New York: Cambridge University Press.

———. 2015. "Information as a Probabilistic Difference Maker." *Australasian Journal of Philosophy* 93, no. 3 (July): 419–443. https://doi.org/10.1080/00048402.2014.993665.

Schank, Jeffrey C., and William C. Wimsatt. 1986. "Generative Entrenchment and Evolution." *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1986:33–60.

Schartner, Michael M., Robin L. Carhart-Harris, Adam B. Barrett, Anil K. Seth, and Suresh D. Muthukumaraswamy. 2017. "Increased Spontaneous MEG Signal Diversity for Psychoactive Doses of Ketamine, LSD and Psilocybin." *Scientific Reports* 7, no. 1 (April): 46421. https://doi.org/10.1038/srep46421.

Schickore, Jutta. 2008. "Doing Science, Writing Science." *Philosophy of Science* 75, no. 3 (July): 323–343. https://doi.org/10.1086/592951.

———. 2020. *About Method: Experimenters, Snake Venom, and the History of Writing Scientifically.* Chicago London: The University of Chicago Press.

Schliesser, Eric. 2019. "Synthetic Philosophy." *Biology & Philosophy* 34, no. 2 (April): 19. https://doi.org/10.1007/s10539-019-9673-3.

———. 2024. "Synthetic Philosophy: A Restatement." *Proceedings of the Aristotelian Society* 124, no. 3 (December): 229–252. https://doi.org/10.1093/arisoc/aoae018.

Schmieder, Falko. 2012. "Georg Toepfers "Historisches Wörterbuch Der Biologie". Zu Methodischen Problemen Einer Begriffsgeschichte Der Biologie." *Archiv für Begriffsgeschichte* 54:267–272.

Schneider, Steven. n.d. "Identity Theory." *Internet Encyclopedia of Philosophy.*

Schnödl, Gottfried, and Florian Sprenger. 2022. *Uexküll's Surroundings: Umwelt Theory and Right-Wing Thought.* 4. Lüneburg: meson press.

Scholl, Raphael, and Tim Räz. 2016. "Towards a Methodology for Integrated History and Philosophy of Science." In *The Philosophy of Historical Case Studies,* edited by Tilman Sauer and Raphael Scholl, 319:69–91. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-30229-4_5.

Schrödinger, Erwin. 1944. *What Is Life? The Physical Aspect of the Living Cell with Mind and Matter and Autobiographical Sketches.* Cambridge ; New York: Cambridge University Press.

Schulte, Peter, and Karen Neander. 2022. "Teleological Theories of Mental Content." In *The Stanford Encyclopedia of Philosophy,* Summer 2022, edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University.

Scott-Phillips, Thom. 2008. "Defining Biological Communication." *Journal of Evolutionary Biology* 21, no. 2 (March): 387–395. https://doi.org/10.1111/j.1420-9101.2007.01497.x.

———. 2025. "Communication & Grammar: A Synthesis." *Psychological Review,* https://doi.org/10.1037/rev0000542.

Scott-Phillips, Thom, and Christophe Heintz. 2023. "Animal Communication in Linguistic and Cognitive Perspective." *Annual Review of Linguistics* 9, no. 1 (January): 93–111. https://doi.org/10.1146/annurev-linguistics-030421-061233.

Scott-Phillips, Thom, Simon Kirby, and Graham R.S. Ritchie. 2009. “Signalling Signalhood and the Emergence of Communication.” *Cognition* 113, no. 2 (November): 226–233. https://doi.org/10.1016/j.cognition.2009.08.009.

Sebeok, Thomas A. 2001. *Signs: An Introduction to Semiotics.* Toronto ; Buffalo: University of Toronto Press.

Segundo-Ortin, Miguel, and Paco Calvo. 2022. “Consciousness and Cognition in Plants.” *WIREs Cognitive Science* 13, no. 2 (March): e1578. https://doi.org/10.1002/wcs.1578.

Seibt, Johanna. 2009. “Forms of Emergent Interaction in General Process Theory.” *Synthese* 166, no. 3 (February): 479–512. https://doi.org/10.1007/s11229-008-9373-z.

Seth, Anil K. 2025. “Conscious Artificial Intelligence and Biological Naturalism.” *Behavioral and Brain Sciences* (April): 1–42. https://doi.org/10.1017/S0140525X25000032.

Seyfarth, Robert M., Dorothy L. Cheney, and Peter Marler. 1980. “Monkey Responses to Three Different Alarm Calls: Evidence of Predator Classification and Semantic Communication.” *Science* 210, no. 4471 (November): 801–803. https://doi.org/10.1126/science.7433999.

Shannon, Claude E. 1948. “A Mathematical Theory of Communication.” *Bell System Technical Journal* 27, no. 3 (July): 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x.

———. (1956) 2021. “The Bandwagon.” In *Information: A Reader,* edited by Eric Hayot, Lea Pao, and Anatoly Detwyler, 51–52. Columbia University Press. https://doi.org/10.7312/hayo18620-007.

Shannon, Claude E., and Warren Weaver. 1964. *The Mathematical Theory of Communication.* Urbana: The University of Illinois Press.

Shea, Nicholas. 2007a. “Consumers Need Information: Supplementing Teleosemantics with an Input Condition.” *Philosophy and Phenomenological Research* 75, no. 2 (September): 404–435. https://doi.org/10.1111/j.1933-1592.2007.00082.x.

———. 2007b. “Representation in the Genome and in Other Inheritance Systems.” *Biology & Philosophy* 22, no. 3 (May): 313–331. https://doi.org/10.1007/s10539-006-9046-6.

———. 2011. “What’s Transmitted? Inherited Information.” *Biology & Philosophy* 26, no. 2 (March): 183–189. https://doi.org/10.1007/s10539-010-9232-4.

———. 2018. *Representation in Cognitive Science.* Oxford, United Kingdom: Oxford University Press.

Shea, Nicholas, Peter Godfrey-Smith, and Rosa Cao. 2018. “Content in Simple Signalling Systems.” *The British Journal for the Philosophy of Science* 69, no. 4 (December): 1009–1035. https://doi.org/10.1093/bjps/axw036.

Shech, Elay, and Axel Gelfert. 2016. *The Exploratory Role of Idealizations and Limiting Cases in Models,* July.

Sheeran, Paschal. 2002. “Intention—Behavior Relations: A Conceptual and Empirical Review.” *European Review of Social Psychology* 12, no. 1 (January): 1–36. https://doi.org/10.1080/14792772143000003.

Shen, Guohua, Tomoyasu Horikawa, Kei Majima, and Yukiyasu Kamitani. 2019. “Deep Image Reconstruction from Human Brain Activity.” Edited by Jill O’Reilly. *PLOS Computational Biology* 15, no. 1 (January): e1006633. https://doi.org/10.1371/journal.pcbi.1006633.

Sims, Matthew. 2021. "A Continuum of Intentionality: Linking the Biogenic and Anthropogenic Approaches to Cognition." *Biology & Philosophy* 36, no. 6 (December). https://doi.org/10.1007/s10539-021-09827-w.

———. 2024. "The Principle of Dynamic Holism: Guiding Methodology for Investigating Cognition in Nonneuronal Organisms." *Philosophy of Science* 91, no. 2 (April): 430–448. https://doi.org/10.1017/psa.2023.104.

———. 2025. *Slime Mould and Philosophy.* Cambridge University Press, January. https://doi.org/10.1017/9781009488648.

Skewes, J. C., and C. A. Hooker. 2009. "Bio-Agency and the Problem of Action." *Biology & Philosophy* 24, no. 3 (June): 283–300. https://doi.org/10.1007/s10539-008-9135-9.

Skyrms, Brian. 2010. *Signals: Evolution, Learning, & Information.* Oxford ; New York: Oxford University Press.

Sloman, Aaron. 2018. *What Did Bateson Mean When He Wrote "Information" Is "a Difference That Makes a Difference"?,* October. Accessed November 29, 2025.

Smaldino, Paul E. 2017. "Models Are Stupid, and We Need More of Them." In *Computational Social Psychology,* 1st ed., edited by Robin R. Vallacher, Stephen J. Read, and Andrzej Nowak, 311–331. New York : Routledge, 2017. | Series: Frontiers of social psychology: Routledge, May. https://doi.org/10.4324/9781315173726-14.

Smith, Daniel W. 2012. "On the Nature of Concepts." *Parallax* 18, no. 1 (February): 62–73. https://doi.org/10.1080/13534645.2012.632976.

Smortchkova, Joulia, Krzysztof Dołęga, and Tobias Schlicht, eds. 2020. *What Are Mental Representations?* New York: Oxford University Press.

Sober, Elliott. 2018. *Philosophy of Biology.* Boulder: Routledge.

Soler, Léna, Sjoerd Zwart, Vincent Israel-Jost, and Michael E. Lynch, eds. 2014. *Science after the Practice Turn in the Philosophy, History, and Social Studies of Science.* 14. New York: Routledge.

Spärck Jones, Karen. 1972. "A Statistical Interpretation of Term Specificity and Its Application in Retrieval." *Journal of Documentation* 28, no. 1 (January): 11–21. https://doi.org/10.1108/eb026526.

Sperber, Dan, and Deirdre Wilson. 1995. *Relevance: Communication and Cognition.* Oxford ; Cambridge, MA: Blackwell Publishers.

Steel, Daniel, Chad Gonnerman, and Michael O'Rourke. 2017. "Scientists' Attitudes on Science and Values: Case Studies and Survey Methods in Philosophy of Science." *Studies in History and Philosophy of Science Part A* 63 (June): 22–30. https://doi.org/10.1016/j.shpsa.2017.04.002.

Steels, Luc. 2015. *The Talking Heads Experiment: Origins of Words and Meanings.* Berlin: Language Science Press. https://doi.org/10.17169/FUDOCS_document_000000022455.

Stegmann, Ulrich E. 2004. "The Arbitrariness of the Genetic Code." *Biology & Philosophy* 19, no. 2 (March): 205–222. https://doi.org/10.1023/B:BIPH.0000024412.82219.a6.

———. 2005. "Genetic Information as Instructional Content." *Philosophy of Science* 72, no. 3 (July): 425–443. https://doi.org/10.1086/498472.

———. 2009. "A Consumer-Based Teleosemantics for Animal Signals." *Philosophy of Science* 76, no. 5 (December): 864–875. https://doi.org/10.1086/605820.

———. 2013. “Introduction: A Primer on Information and Influence in Animal Communication.” In *Animal Communication Theory: Information and Influence,* edited by Ulrich E. Stegmann, 1–40. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139003551.002.

———. 2014. “Causal Control and Genetic Causation.” *Noûs* 48, no. 3 (September): 450–465. https://doi.org/10.1111/j.1468-0068.2012.00867.x.

———. 2016. “The Philosophy of Communication and Information.” In *The Routledge Handbook of Philosophy of Information,* edited by Luciano Floridi. Routledge Handbooks in Philosophy. London ; New York: Routledge, Taylor & Francis Group.

Sterelny, Kim. 2000. “The “Genetic Program” Program: A Commentary on Maynard Smith on Information in Biology.” *Philosophy of Science* 67, no. 2 (June): 195–201. https://doi.org/10.1086/392769.

———. 2010. “Minds: Extended or Scaffolded?” *Phenomenology and the Cognitive Sciences* 9, no. 4 (December): 465–481. https://doi.org/10.1007/s11097-010-9174-y.

Stotz, Karola, and Colin Allen. 2012. “From Cell-Surface Receptors to Higher Learning: A Whole World of Experience.” In *Philosophy of Behavioral Biology,* edited by Kathryn S. Plaisance and Thomas A.C. Reydon, 282:85–123. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-1951-4_5.

Stotz, Karola, and Paul E. Griffiths. 2017. “Biological Information, Causality and Specificity - an Intimate Relationship.” In *From Matter to Life: Information and Causality,* edited by Sara Imari Walker, Paul C. W. Davies, and George F. R. Ellis. Cambridge University Press.

Strevens, Michael. 2013. “No Understanding without Explanation.” *Studies in History and Philosophy of Science Part A* 44, no. 3 (September): 510–515. https://doi.org/10.1016/j.shpsa.2012.12.005.

Suárez, Javier, and Adrian Stencel. 2020. “A Part-dependent Account of Biological Individuality: Why Holobionts Are Individuals *and* Ecosystems Simultaneously.” *Biological Reviews* 95, no. 5 (October): 1308–1324. https://doi.org/10.1111/brv.12610.

Suárez, Mauricio, ed. 2008. *Fictions in Science: Philosophical Essays on Modeling and Idealization.* Routledge, October. https://doi.org/10.4324/9780203890103.

Suppe, Frederick. 1998. “The Structure of a Scientific Paper.” *Philosophy of Science* 65 (3): 381–405.

Swammerdam, Jan. (1676–79) 1758. *The Book of Nature.* Edited by John Hill. Translated by Thomas Flloyd. London: Seyffert.

Tabachnick, Barbara G., and Linda S. Fidell. 2014. *Using Multivariate Statistics.* Harlow, Essex, England: Pearson.

Thagard, Paul. 1988. *Computational Philosophy of Science.* Cambridge, Mass: MIT Press.

———. 2003. “Pathways to Biomedical Discovery.” *Philosophy of Science* 70, no. 2 (April): 235–254. https://doi.org/10.1086/375465.

Thelen, Esther, and Linda B. Smith. 1994. *A Dynamic Systems Approach to the Development of Cognition and Action.* The MIT Press, July. https://doi.org/10.7551/mitpress/2524.001.0001.

Thomasson, Amie. 2015. “What Can Philosophy Really Do?” *The Philosophers’ Magazine,* no. 71, 17–23. https://doi.org/10.5840/tpm20157196.

Thomasson, Amie L. 2017. "Metaphysics and Conceptual Negotiation." *Philosophical Issues* 27, no. 1 (October): 364–382. https://doi.org/10.1111/phis.12106.

Thompson, Evan. 2010. *Mind in Life: Biology, Phenomenology, and the Sciences of Mind.* Cambridge, Mass.: Belknap Press of Harvard Univ. Press.

Thomson, Spencer, and Henry H. Smith. 1853. *A Dictionary of Domestic Medicine and Household Surgery.* Philadelphia: Lippincott, Grambo & co.

Tilton, Lauren, David Mimno, and Jessica Marie Johnson, eds. 2024. *Computational Humanities.* Minneapolis, MN: University of Minnesota Press.

Tinbergen, N. 1963. "On Aims and Methods of Ethology." *Zeitschrift für Tierpsychologie* 20, no. 4 (January): 410–433. https://doi.org/10.1111/j.1439-0310.1963.tb01161.x.

Toepfer, Georg. 2011. *Historisches Wörterbuch der Biologie: Geschichte und Theorie der biologischen Grundbegriffe. Band 2: Gefühl — Organismus.* Stuttgart: J.B. Metzler. https://doi.org/10.1007/978-3-476-00455-0.

———. 2024. *BioConcepts: The Origin and Definition of Biological Concepts - A Multilingual Database.* Accessed November 15, 2024.

Tolman, Edward C. 1948. "Cognitive Maps in Rats and Men." *Psychological Review* 55 (4): 189–208. https://doi.org/10.1037/h0061626.

Triandis, Harry C. 1980. "Values, Attitudes, and Interpersonal Behavior." In *Nebraska Symposium on Motivation 1979,* edited by Monte M. Page, 195–260. Lincoln/London: University of Nebraska Press.

Turkheimer, Federico E., Fernando E. Rosas, Ottavia Dipasquale, Daniel Martins, Erik D. Fagerholm, Paul Expert, František Váša, Louis-David Lord, and Robert Leech. 2022. "A Complex Systems Perspective on Neuroimaging Studies of Behavior and Its Disorders." *The Neuroscientist* 28, no. 4 (August): 382–399. https://doi.org/10.1177/1073858421994784.

Ugbebor, Olabisi O. 1980. "Ergodic Properties of Brownian Motion." *Proceedings of the Edinburgh Mathematical Society* 23, no. 3 (October): 331–340. https://doi.org/10.1017/S0013091500003874.

Underwood, Ted, and Jordan Sellars. 2012. "The Emergence of Literary Diction." *Journal of Digital Humanities* 1 (2).

Van Gelder, Tim. 1995. "What Might Cognition Be, If Not Computation?:" *Journal of Philosophy* 92 (7): 345–381. https://doi.org/10.2307/2941061.

———. 1998. "The Dynamical Hypothesis in Cognitive Science." *Behavioral and Brain Sciences* 21, no. 5 (October): 615–628. https://doi.org/10.1017/S0140525X98001733.

Vandenberg, Laura N., Dany S. Adams, and Michael Levin. 2012. "Normalized Shape and Location of Perturbed Craniofacial Structures in the Xenopus Tadpole Reveal an Innate Ability to Achieve Correct Morphology." *Developmental Dynamics* 241, no. 5 (May): 863–878. https://doi.org/10.1002/dvdy.23770.

Vanderwolf, C. H. 2007. *The Evolving Brain: The Mind and the Neural Control of Behavior.* New York, NY: Springer.

Varela, Francisco J. 1986. "Steps to a Cybernetics of Autonomy." In *Power, Autonomy, Utopia: New Approaches Toward Complex Systems,* edited by Robert Trappl, 117–122. Boston, MA: Springer US. https://doi.org/10.1007/978-1-4613-2225-2_8.

Varela, Francisco J., Evan Thompson, and Eleanor Rosch. 1991. *The Embodied Mind: Cognitive Science and Human Experience.* Cambridge, Mass: MIT Press.

Varga, Somogy. 2021. "The Continuity of Inquiry and Normative Philosophy of Science." *Metaphilosophy* 52, no. 5 (October): 655–667. https://doi.org/10.1111/meta.12513.

Veigl, Sophie Juliane. 2022. "Do Heritable Immune Responses Extend Physiological Individuality?" *History and Philosophy of the Life Sciences* 44, no. 4 (December): 67. https://doi.org/10.1007/s40656-022-00549-0.

Verspoor, Karin, K Bretonnel Cohen, and Lawrence Hunter. 2009. "The Textual Characteristics of Traditional and Open Access Scientific Journals Are Similar." *BMC Bioinformatics* 10, no. 1 (December): 183. https://doi.org/10.1186/1471-2105-10-183.

Vert, Grégory, and Joanne Chory. 2011. "Crosstalk in Cellular Signaling: Background Noise or the Real Thing?" *Developmental Cell* 21, no. 6 (December): 985–991. https://doi.org/10.1016/j.devcel.2011.11.006.

Volterra, Vito. 1931. "Variations and Fluctuations of the Number of Individuals in Animal Species Living Together." In *Animal Ecology,* edited by R.N. Chapman, translated by Mary Evelyn Wells, 409–448. New York: McGraw-Hill.

Von Neumann, John. 1966. *Theory of Self-Reproducing Automata.* Edited by Arthur W. Burks. Urbana; London: The University of Illinois Press.

Von Neumann, John, and Oskar Morgenstern. (1944) 2007. *Theory of Games and Economic Behavior.* Princeton, NJ: Princeton University Press.

Walton, Kendall L. 1993. *Mimesis as Make-Believe: On the Foundations of the Representational Arts.* Cambridge, Mass.: Harvard University Press.

Wan, Kirsty Y., and Gáspár Jékely. 2021. "Origins of Eukaryotic Excitability." *Philosophical Transactions of the Royal Society B: Biological Sciences* 376, no. 1820 (March): 20190758. https://doi.org/10.1098/rstb.2019.0758.

Wang, Jianguo, Xiaomeng Yi, Rentong Guo, Hai Jin, Peng Xu, Shengjun Li, Xiangyu Wang, et al. 2021. "Milvus: A Purpose-Built Vector Data Management System." In *Proceedings of the 2021 International Conference on Management of Data,* 2614–2627. Virtual Event China: ACM, June. https://doi.org/10.1145/3448016.3457550.

Washburn, Margaret Floy. 1908. *The Animal Mind: A Text-book of Comparative Psychology.* New York: The Macmillan Company.

Watson, J. D., and F. H. C. Crick. 1953. "Genetical Implications of the Structure of Deoxyribonucleic Acid." *Nature* 171, no. 4361 (May): 964–967. https://doi.org/10.1038/171964b0.

Weber, Marcel. 2005. "Genes, Causation and Intentionality." *History and Philosophy of the Life Sciences* 27 (3/4): 407–420.

Weisberg, Michael. 2007. "Three Kinds of Idealization:" *Journal of Philosophy* 104 (12): 639–659. https://doi.org/10.5840/jphil20071041240.

———. 2013. *Simulation and Similarity: Using Models to Understand the World.* New York: Oxford University Press.

Werning, Markus, and Sen Cheng. 2017. "Taxonomy and Unity of Memory." In *The Routledge Handbook of Philosophy of Memory,* First edition, edited by Sven Bernecker and Kourken Michaelian. Routledge Handbooks in Philosophy. London: Taylor and Francis. https://doi.org/10.4324/9781315687315.

Weron, Aleksander, Krzysztof Burnecki, Elizabeth J. Akin, Laura Solé, Michał Balcerek, Michael M. Tamkun, and Diego Krapf. 2017. "Ergodicity Breaking on the Neuronal Surface Emerges from Random Switching between Diffusive States." *Scientific Reports* 7, no. 1 (July): 5404. https://doi.org/10.1038/s41598-017-05911-y.

Whitehead, Alfred North. (1927) 1985. *Symbolism: Its Meaning and Effect.* New York, NY: Fordham.

WHO. *Diabetes.* Accessed May 21, 2025.

Wieling, Martijn, Josine Rawee, and Gertjan Van Noord. 2018. "Reproducibility in Computational Linguistics: Are We Willing to Share?" *Computational Linguistics* 44, no. 4 (December): 641–649. https://doi.org/10.1162/coli_a_00330.

Wiener, Norbert. (1948) 2019. *Cybernetics: Or, Control and Communication in the Animal and the Machine.* Cambridge, MA: The MIT Press.

Wiley, R Haven. 1994. "Errors, Exaggeration, and Deception in Animal Communication." Edited by Leslie Real. *Behavioral mechanisms in evolutionary ecology* 161:157–189.

———. 2013. "Communication as a Transfer of Information: Measurement, Mechanism and Meaning." In *Animal Communication Theory: Information and Influence,* edited by Ulrich E. Stegmann, 113–131. Cambridge ; New York: Cambridge University Press.

Wilkenfeld, Daniel A. 2019. "Understanding as Compression." *Philosophical Studies* 176, no. 10 (October): 2807–2831. https://doi.org/10.1007/s11098-018-1152-1.

Wilson, Deirdre, and Dan Sperber. 2010. "Relevance Theory." In *The Handbook of Pragmatics,* 5. [print.], edited by Laurence R. Horn. Blackwell Handbooks in Linguistics 16. Malden, Mass.: Blackwell.

Wilson, Edward O. 1975. *Sociobiology: The New Synthesis.* Cambridge, Mass.: Belknap Pr. of Harvard Univ. Pr.

Wilson, Margaret. 2002. "Six Views of Embodied Cognition." *Psychonomic Bulletin & Review* 9, no. 4 (December): 625–636. https://doi.org/10.3758/BF03196322.

Wilson, Mark. 2006. *Wandering Significance: An Essay on Conceptual Behavior.* Oxford: Clarendon Press.

Wimsatt, William C. 2007. *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality.* Cambridge, Mass: Harvard University Press.

Winkelbauer, Karel. 1970. "On the Asymptotic Rate of Non-Ergodic Information Sources." *Kybernetika* 6 (2): 127–148.

Winsberg, Eric. 2018. *Philosophy and Climate Science.* Cambridge: Cambridge University Press.

Winther, Rasmus Grønfeldt. 2021. "The Structure of Scientific Theories." In *The Stanford Encyclopedia of Philosophy,* Spring 2021, edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University.

Wittgenstein, Ludwig. 1968. *Philosophical investigations.* Translated by G. E. M. Anscombe. Oxford: Basil Blackwell.

Witzany, Guenther. 2019. "Communication as the Main Characteristic of Life." In *Handbook of Astrobiology,* edited by Vera M. Kolb, 91–105. Boca Raton, FL: CRC Press.

———. 2020. "What Is Life?" *Frontiers in Astronomy and Space Sciences* 7 (March): 7. https://doi.org/10.3389/fspas.2020.00007.

Woodward, James. 2004. *Making Things Happen: A Theory of Causal Explanation.* Oxford University PressNew York, January. https://doi.org/10.1093/0195155270.001.0001.

———. 2010. "Causation in Biology: Stability, Specificity, and the Choice of Levels of Explanation." *Biology & Philosophy* 25, no. 3 (June): 287–318. https://doi.org/10.1007/s10539-010-9200-z.

———. 2023. "Causation and Manipulability." In *The Stanford Encyclopedia of Philosophy,* Summer 2023, edited by Edward N. Zalta and Uri Nodelman. Metaphysics Research Lab, Stanford University.

Yousif, Abdallah, Zhendong Niu, James Chambua, and Zahid Younas Khan. 2019. "Multi-Task Learning Model Based on Recurrent Convolutional Neural Networks for Citation Sentiment and Purpose Classification." *Neurocomputing* 335 (March): 195–205. https://doi.org/10.1016/j.neucom.2019.01.021.

Yuste, Rafael, Rosa Cossart, and Emre Yaksi. 2024. "Neuronal Ensembles: Building Blocks of Neural Circuits." *Neuron* 112, no. 6 (March): 875–892. https://doi.org/10.1016/j.neuron.2023.12.008.

Zhang, Jiaje, and Donald A. Norman. 1994. "Representations in Distributed Cognitive Tasks." *Cognitive Science* 18, no. 1 (January): 87–122. https://doi.org/10.1207/s15516709cog1801_3.

Zhou, Binghan, Wanling Lin, Yaling Long, Yunkai Yang, Huan Zhang, Kongming Wu, and Qian Chu. 2022. "Notch Signaling Pathway: Architecture, Disease, and Therapeutics." *Signal Transduction and Targeted Therapy* 7, no. 1 (March): 95. https://doi.org/10.1038/s41392-022-00934-y.

Zollman, Kevin J. S., Carl T. Bergstrom, and Simon M. Huttegger. 2013. "Between Cheap and Costly Signals: The Evolution of Partially Honest Communication." *Proceedings of the Royal Society B: Biological Sciences* 280, no. 1750 (January): 20121878. https://doi.org/10.1098/rspb.2012.1878.

Zubek, Julian, Tomasz Korbak, and Joanna Rączaszek-Leonardi. 2024. "Models of Symbol Emergence in Communication: A Conceptual Review and a Guide for Avoiding Local Minima." *Artificial Intelligence Review* 58, no. 2 (December): 65. https://doi.org/10.1007/s10462-024-11048-y.