

# Carnap, Feyerabend and the pragmatic theory of observation

Daniel Kuby

## Abstract

Paul Feyerabend once made a remark to the effect that his pragmatic theory of observation can be traced back to proposals put forward by leading Logical Empiricists during the height of the protocol sentence debate. In this paper I want to vindicate the systematic side of Feyerabend's remark and show that a pragmatic theory of observation can in fact be found in Rudolf Carnap's writings of 1932. I first proceed to dispel a misunderstanding concerning the term "pragmatic" raised by Thomas Oberdan. Following Morris' and Carnap's documented usage, I show that the intended meaning of "pragmatic" refers to a specific semiotic relation between users of a language and their environment describable by empirical means (specifically, the causal relation between the verbal behavior of language users and their environment). I reconstruct such a pragmatic theory in terms of a detector model that interprets observation sentences as bodily dispositions indicating physical events in the surroundings of the detector. I then proceed to show how Feyerabend's later theory of observation picks up central features of Carnap's account and also shares some of the motivations. I conclude by noting how an empirical theory of observation sentences offers a bootstrapping solution to the "basis problem".

# Acknowledgements

The author would like to thank Matteo Collodel, Björn Henning, Tom Oberdan, as well as Christian Damböck, Richard Dawid, Christoph Limbeck-Lilienau and other attendees of the Institute Vienna Circle Colloquium for valuable feedback on earlier drafts of this paper. This paper first appeared as a chapter in my Ph.D thesis (2017) and I am grateful to my dissertation advisors, Elisabeth Nemeth and Wolfgang Reiter, as well as to the external examiners, Don Howard and Thomas Uebel, for their critical comments. This work has been funded by the Austrian Science Fund (FWF W1228-G18 grant).

## 1. Introduction

In the late 1950s and early 1960s, Paul Feyerabend made waves in the philosophy of science community by advancing a number of sweeping proposals, among them a radical version of the theory-ladenness of observation and the incommensurability-thesis (Feyerabend 1958; Feyerabend 1962). While theory-ladenness and incommensurability were hotly debated by other contemporaries and became stock in the history of philosophy of science (not least because some versions were advocated by other notable philosophers (Hanson 1958; Kuhn 1962) and could be traced back in history (Duhem 1906)), a third incisive proposal by Feyerabend, the so-called Pragmatic Theory of Observation (PTO) first published in his (1958), has not become a classical topic. Various explanations are readily available: seemingly no other ‘post-positivist’ philosopher of science advocated it and it has been discussed mostly dismissively ever since (Butts 1966; Townsend 1970; Hull 1972).<sup>1</sup> Also, the conceptual resources used to articulate PTO are chiefly rooted in the statement view. While incom-

---

<sup>1</sup>Important exceptions are Grover Maxwell’s classic paper “On the ontological status of theoretical entities”, in which PTO appeared briefly as the conception of “quickly decidable sentences” (Maxwell 1962, 13); and the early Bas van Fraassen (cf. his (1992)), who inherited a cognate account from Wilfrid Sellars and was sympathetic to PTO. (This was confirmed to me by van Fraassen in personal communication.) A clear account of PTO in Feyerabend scholarship is given by Oberheim (2006).

measurability and theory-ladenness could be discussed even after its abandonment, PTO apparently could not.

Though PTO has not been passed down ‘internally’, it has been introduced into the historical discussion by Oberdan (1990), who evaluates PTO in a historical perspective and is cited ever since (see for example Preston 1997, 45–50, 214; Oberheim 2006, 233; Misak 1995, 80, 202; Olegario da Silva, Gilson 2013, 155–6). In this paper I want to revive the conversation in the HOPOS community. The setting of this discussion is quite interesting. In his (1962), Feyerabend not only advances PTO, but also sketches a genealogy, tracing the idea of a “pragmatic” theory of observation back to Logical Empiricism—a notable circumstance, given the very critical attitude with regard to the received view of Logical Empiricism for which he is generally known. However, from the vantage point of an historical appraisal of Feyerabend’s formative years, this claim doesn’t come as a surprise. The background thesis of the present paper, not developed here any further, is that Feyerabend’s philosophical coming-of-age took place within the context of scientific philosophy—and Logical Empiricism in particular.<sup>2</sup> In this sense, a historically sound interpretation of Feyerabend’s philosophy and its development can only be had by taking into account our best knowledge of the history of scientific philosophy.<sup>3</sup>

Furthermore, Feyerabend’s claim doubles his role: PTO makes him an historical protagonist and, at the same time, he becomes an historical commentator of PTO. It is the latter role that has been taken up by Oberdan (1990), who advances a critique of Feyerabend’s historical

---

<sup>2</sup>See Stadler (2010), Kuby (2010), also Yuann (2007). The most detailed missing link, Feyerabend’s unpublished dissertation *Zur Theorie der Basissätze* (1951), is still unpublished, but other textual evidence has surfaced in the meantime: Feyerabend ([1951] 2010), Feyerabend ([1948] 2016); see Kuby, Limbeck-Lilienau, and Schorner (2010), Kuby (2016) for commentary.

<sup>3</sup>Logical Empiricism is but one of many attempts at the creation of a “scientific philosophy”. The term is used here to denote a philosophy which self-identifies with an orientation towards the sciences. The question how this orientation should be understood and the kind of “scientification” of philosophy it engenders has been answered very differently by different strains within the tradition of scientific philosophy (see Richardson 1997).

contention that the development of PTO can be traced back to Logical Empiricism, especially to Rudolf Carnap. In his critique, Oberdan claims that Feyerabend misunderstood Carnap's view on protocol sentences, and, more specifically, that the latter never held the particular characterization of observation sentences that Feyerabend ascribes to him, to wit, a "pragmatic" account of observation sentences.<sup>4</sup> Contrary to this reading, I advance the claim that Feyerabend's actual usage of the expression "pragmatic" matches an interesting proposal about protocol sentences emerging in the protocol sentence debate of the early 1930s, particularly in Carnap's writings.

If an historical appreciation of Logical Empiricism is necessary to interpret Feyerabend's philosophy, conversely, I think that Feyerabend can help the systematic study of Logical Empiricism in HOPOS scholarship. The present paper is a first attempt at a vindication of this claim, focusing on a neglected result of the Vienna Circle's discussions, and reconstructing a pragmatic characterization of observation as can be found in Carnap's writings of 1932.

I will proceed as follows: I first argue that Oberdan's critique misses the mark, as it misreads Feyerabend's use of "pragmatic" in the context of PTO and, as a consequence, incorrectly tracks the genealogy claimed by Feyerabend. I argue that Feyerabend's actual usage of the expression "pragmatic"—following Morris' definition, and as synonym for "causal"—gives a more adequate interpretation of what PTO is about. I argue that Feyerabend refers

---

<sup>4</sup>Oberdan makes a further, more interesting objection, which cannot be addressed here due to limitations of space. According to Oberdan, Feyerabend's characterization of observation sentences has a *universal* pretension, which is alien to Carnap's language-relative outlook: "For Feyerabend specified pragmatic (including causal) criteria for observation reports that are independent of any interpretation they might receive in a given language. Then the criterion of observability transcends the limits of any given theoretical framework, yielding universal criteria for distinguishing observation reports. Of course, the very notion of universal criteria [...] is alien to Carnap's conception" (Oberdan 1990, 30–31). A brief comment: since Feyerabend's causal account of observation sentences closely matches Carnap's account in this regard, either such a universal pretension in Carnap's account has gone unnoticed (after all, the account precedes his adoption of the tolerance principle)—or the kind of universal pretension ascribed to Feyerabend can't possibly be correct. I tend to the latter.

to an overlooked empirical characterization of protocol sentences to be found in Carnap's writings of 1932. Noting the traditional view of Carnap as interested chiefly in logical matters, I contextualize this synthetic, empirical account within the protocol sentence debate (§2). In the subsequent section I give a detailed reconstruction of such an empirical characterization of protocol sentences as can be found in Carnap's *Erkenntnis* articles of 1932, working out what I call the detector model of observation (or protocol) sentences (§3). I then go back to sketch some of the most obvious lessons Feyerabend adopted from Carnap's proposal (§4) and conclude my exposition by commenting on the pragmatic account of observation sentences as a solution to the "basis problem" (§5).

## 2. Setting the stage: An empirical account of the empirical basis of science

### 2.1 Feyerabend's and Carnap's notion of "pragmatic"

Let's first introduce Oberdan's account of PTO, according to which PTO is composed of "two ideas": "The first is that observation reports are 'theory-laden' in the sense that they are always interpreted in the light of the best current theory and are subject to reinterpretation when one theory succeeds another"<sup>5</sup> (Oberdan 1990, 25).

According to Oberdan, "Feyerabend finds that the theory-ladenness of observation leads inexorably to the second idea of the Pragmatic Theory", namely, that "what counts as an observation report does not depend on either its empirical content or its logical form, but on its causal or pragmatic features" (Oberdan 1990, 25).<sup>6</sup>

---

<sup>5</sup>Though Oberdan does not give any references to Feyerabend's work, this is clearly a paraphrase of Feyerabend's "thesis I" first presented in 1958: [...] *[T]he interpretation of an observation language is determined by the theories that we use to explain what we observe, and it changes as soon as those theories change.* (Feyerabend 1958, 163)

<sup>6</sup>Notably, Oberdan agrees with the—in my view much more contentious—claim that the "first idea"

For the sake of argument, we will work with Oberdan's summary of PTO and, by doing so, overlook the problematic nature of this account, in order to address his main contention: "[T]he function Carnap allotted to pragmatic factors in the discrimination of observation reports is nothing like what Feyerabend has in mind. Then it cannot be said that Carnap endorsed the second idea of the Pragmatic Theory in anything like the sense in which Feyerabend intended. These considerations suffice to show that Feyerabend's historical contentions about Carnap's early philosophy are false" (Oberdan 1990, 25).

What did Carnap have in mind that Feyerabend did not?

Oberdan interprets the expression "pragmatic" as referring to Carnap's famous conventionalized understanding of language expressed in his Principle of Tolerance.<sup>7</sup> In Oberdan's words: "The choice of a language is independent of objective matters of fact so that only *pragmatic* considerations guide the decision to adopt a given form of language" (Oberdan 1990, 29, emphasis added). Oberdan then tries to give a charitable reading of the work that the Principle can do for a "pragmatic" theory of observation: the application of the Principle to the problem of characterizing the observation language yields the conclusion that "there is no absolute class of observational sentences, fixed once and for all time [...]. Rather, distinguishing certain statements as observational is part of determining the specifics of a language. Since these matters are conventional, they ultimately turn on *pragmatic* considerations and, in this limited sense, Carnap's account may be termed 'pragmatic'" (Oberdan 1990, 29, emphasis added).

But even admitting this, Oberdan argues, a *pragmatic* theory of observation sentences, as Feyerabend would have it, does not follow. More specifically, Oberdan's first charge

---

of Feyerabend's PTO, the theory-ladenness of observation as expressed in thesis I, can indeed be found in Carnap as a natural consequence of his Metalogic Thesis: "Since the meaning of statements are given by their inferential relations to other statements, this last conclusion [i.e. that two sentences are regarded as inferentially related whenever current theory vindicates the conclusion of one from the other] is tantamount to the Pragmatic Theory's notion that observation reports are to be translated in light of contemporary scientific thought" (Oberdan 1990, 27).

<sup>7</sup>See Carnap (1934a, sec. 17).

against Feyerabend is one of omission: Feyerabend negligently ignores the linguistic background of Carnap's "pragmatism", thus misconstruing Carnap's conventionalism.

A very basic observation shows Oberdan's charge to be misplaced: Feyerabend's use of "pragmatic" doesn't refer at all to Carnap's conventionalism. Feyerabend's theory is not "pragmatic" in the sense that it refers to external framework questions.<sup>8</sup> Anticipating my result in Carnap's parlance: Feyerabend's PTO is a theoretical proposal. The "pragmatic" component in it *doesn't refer to the external determination of the specifics of a language, but to a specific result of such a determination*. In other words, Feyerabend's genealogical claim doesn't refer to Carnap's contributions to metaphilosophy, but to the content of one of Carnap's early language proposals.

Feyerabend makes his understanding of "pragmatic" explicit in the very paper Oberdan criticizes. Commenting on passages by Popper and Carnap allegedly indicating their 'retreat from experience', Feyerabend introduces the terminology thus: "Now it is most important to realize that the characterization of observation statements implicit in the above quotations is a *causal* characterization, or if one wants to use more recent terminology, a *pragmatic* characterization" (Feyerabend 1962, 36). The first glaring hint is that "pragmatic" is introduced as synonym for "causal" (not "conventional"). Further, Feyerabend introduces the word "pragmatic" to conform to a "more recent terminology", which is spelled out in the footnote, where Morris' *Foundations of the Theory of Signs* (1938), in particular his definition of "pragmatics", is referenced.

Famously, "pragmatics" is a technical term introducing a third "dimension of semiosis" (Morris 1938, 30) alongside syntax and semantics: "By 'pragmatics' is designated the sci-

---

<sup>8</sup>This is not to say that PTO is not *also* "pragmatic" in Oberdan's sense. In fact, Feyerabend claims explicitly that the epistemic status of PTO is that of a convention: "The choice between the pragmatic theory [of observation] and the semantic theory [of observation] is of course purely a matter of convention" (Feyerabend 1965b, 243). It is a question of choice, not truth. But this further claim, the conventional nature of PTO, is not the reason Feyerabend labels his theory of observation as "pragmatic". (To be sure, Feyerabend's understanding of "choice" differs from Carnap's in that it involves axiological considerations that seem to go beyond and above Carnap's limited instrumental considerations. But this is a topic for another paper.)

ence of the relation of signs to their interpreters. ‘Pragmatics’ must then be distinguished from ‘pragmatism,’ and ‘pragmatical’ from ‘pragmatic.’ Since most, if not all, signs have as their interpreters living organisms, it is a sufficiently accurate characterization of pragmatics to say that it deals with the biotic aspects of semiosis, that is, with all the psychological, biological, and sociological phenomena which occur in the functioning of signs” (Morris 1938, 30).

Pragmatics and cognates were not always used by Morris to separate the realm of pragmatics from syntax and semantics. Sometimes pragmatics denotes the whole field of semiotics. On other occasions pragmatics has a more narrow meaning, denoting only the linguistic sub-discipline.<sup>9</sup> Feyerabend’s characterization strongly suggests that he uses it in the first, differentiating sense. In this limited sense Morris can help us to understand Feyerabend’s notion of pragmatics, but there’s no evidence that Feyerabend’s terminological choice implies a commitment to Morris’s overall semiotic program.

Not so in Carnap’s case. Morris’ tripartition was proposed as a systematic framework for the *International Encyclopedia of Unified Science* and Carnap adopted it quite readily in his own contribution (Carnap [1939] 1957). As defined by Carnap, “pragmatics is an empirical discipline dealing with a special kind of human behavior and making use of the results of the different branches of science (principally social science, but also physics, biology and psychology)” (Carnap [1939] 1957, 6). Pragmatics is concerned with “the action, state, and environment of a man who speaks or hears” (Carnap [1939] 1957, 4). It studies “the mode of use” of expressions and “both the cause and the effect” of utterances in social contexts.

It is most important to notice that the field of pragmatics is in no way coincident with the Carnapian realm of external questions.<sup>10</sup> To the contrary, the questions addressed in pragmatics are theoretical, not practical. Carnap’s life-long dichotomy between theoretical and practi-

---

<sup>9</sup>See Uebel (2013) on this point.

<sup>10</sup>This is true for Carnap’s understanding. As Uebel (2013, 530–1) notes, this does not hold for the very wide meaning of pragmatics occasionally used by Morris, which included axiological considerations, moral-political discourse as ‘humanistic implications of semiotics’.



cal questions prevented him from ever equating “pragmatical” and “practical”. Pragmatical questions are internal questions of the discipline of pragmatics asking for a descriptive answer; practical questions are external questions in need of a volitional decision.<sup>11</sup> Thus, when Carnap characterizes something as “pragmatical”, he refers to its empirical character, not to its conventional or volitional nature.<sup>12</sup>

The same holds for Feyerabend’s use of “pragmatic”. Applied to the characterization of observation sentences, “pragmatic” advertises the circumstance that the observational character is spelled out in terms of the relation between observation sentences (a class of signs) and observers (a class of interpreters), as opposed to the formal relations between sentences (syntax), or the relations between observation sentences and what is observed (semantics). Feyerabend’s earlier designation—“pragmatic” is introduced as a synonym for “causal”—specifies the kind of relation that PTO establishes between observation sentences and observers: a causal relation.

Oberdan’s second charge is that Feyerabend’s PTO cannot possibly be a consequence of Carnap’s conventionalism:

If Carnap’s account of observation is pragmatic, its pragmatism is nonetheless conditioned by the principles of his philosophy of language. If the conventional choice of a language is broadly construed as the choice of a framework of justification, and only pragmatic considerations can influence this choice,

---

<sup>11</sup>See also Uebel (2013), p. 540 on this point.

<sup>12</sup>Complicating the interpretation somewhat, Morris stipulates the adjectival form to be used when referring to the field on inquiry (“pragmatical”), in order to distinguish it from features pertaining to practical considerations (“pragmatic”). When denoting the former, Carnap followed Morris’ terminology down to this specific recommendation, while Feyerabend did not. (Doing so would have probably avoided the misunderstanding present in Oberdan’s paper.) On the other hand, the “pragmatical”-“pragmatic” distinction is not readily available in German: where Carnap used “pragmatisch” in *Logische Syntax*, he himself suggested “pragmatical” (first choice) or “pragmatic” (second choice) in the English translation (cf. Quine and Carnap (1991), p. 275). And, contrary to Morris, he consistently continued to render “praktisch” as “practical”, not “pragmatic”.

then the decision to count certain sentences as observational is likewise a pragmatic issue. But Feyerabend attributes Carnap the *further* conclusion that observation reports are not distinguished by their contents, but by their causes; certain utterances are observation reports because their etiology conforms to certain patterns of behavior. Thus Feyerabend *further* misconstrues Carnap's conventionalism with respect to protocols as if it implied that "an observation sentence is distinguished from other sentences of a theory not by its content but by the cause of its production, by the fact that its production conforms to certain behavioral patterns." (Feyerabend 1962, 36) But this last notion, the idea that observation reports are pragmatically distinguished in the sense that their demarcation depends on their causal features, is a far cry from what Carnap had in mind when he granted pragmatic factors a role in the isolation of observation reports from other scientific statements. (Oberdan 1990, 30, emphasis mine)

I agree with Oberdan that the two notions are "a far cry" from each other. But neither the following charge holds: Feyerabend does not *misconstrue* Carnap's conventionalism and he does not claim that it *implies* PTO, because he does not refer to Carnap's framework conventionalism in the first place—and thus cannot possibly claim the inference that Oberdan attributes to him. PTO is not a "further conclusion" of Carnap's framework conventionalism. It is—in Carnap's parlance—a particular proposal. And this proposal, so Feyerabend claims, Carnap (first) put forward.<sup>13</sup> Having established a charitable interpretation of Feyerabend's notion of a "pragmatic" theory of observation, it is this claim that ought to be carefully checked.<sup>14</sup>

---

<sup>13</sup>In chapter II of Oberdan (1993), in particular the section "The pragmatics of protocols", Oberdan gives a clear account of Carnap's Metalogic and how this conception yielded a pragmatic account of protocol sentences, which makes his misunderstanding of Feyerabend the more puzzling.

<sup>14</sup>Oberdan frames the correctness of this claim as crucial to Feyerabend's own philosophical aims: [A]lthough the purpose behind Feyerabend's explanations is to discredit the later Positivist account of ob-

## 2.2 Feyerabend on the development of empiricism in the 1930s

In his (1962), Feyerabend sets the task of “free[ing]” contemporary empiricism from “the elements which it still shares with its more dogmatic opponents”, paving the way for a “disinfected empiricism”. But, in an unexpected narrative turn, he warns us that this isn’t an original achievement for already “the empiricism of the thirties was disinfected in the sense desired here” (Feyerabend 1962, 31). Judging from his subsequent references, it is evident that he refers to Viennese Logical Empiricism and (what today is known as) the protocol sentence debate.<sup>15</sup> Feyerabend sketches a radical development of empiricism from the “earlier positivism of the Vienna Circle” (Feyerabend 1962, 31) leading up to the debate: “[S]ince these happy and carefree days of the *Aufbau*, logical empiricism has been greatly modified. The changes that took place were mainly of two kinds. On the one side, new ideas were introduced concerning the relation between observational terms and theoretical terms. On the other side, the assumptions made about the observational language itself were modified. In both cases the changes were quite drastic” (Feyerabend 1962, 34–35).<sup>16</sup>

servation, the ultimate success of [Feyerabend’s] systematic program depends on the historical thesis that Carnap endorsed the fundamental tenet of the Pragmatic Theory in his Thirties’ writings on protocols (Oberdan 1990, 25). This seems to me a non-sequitur. The “ultimate success” of Feyerabend’s *systematic program* is not dependent on what Carnap endorsed at a particular time or not. Suppose e.g. that the “ultimate success” of Feyerabend’s systematic program rests on the consistency of PTO; surely the consistency of PTO is not dependent on whether someone else held it or not. It may be the case that Feyerabend’s *critique* of (the received view of) Logical Empiricism—only where it takes the form of a *reductio ad absurdum* argument—is dependent upon the premises actually held by the argumentative opponent, but certainly any such *reductio ad absurdum* is not identifiable with Feyerabend’s systematic program (at least for any reasonable definition of “systematic program”, i.e. entailing the advance of some affirmative claims).

<sup>15</sup>See Oberdan (1993), Uebel (2007) for detailed historical and systematic reconstructions of the protocol sentence debate.

<sup>16</sup>In this development Feyerabend (1962, 35–36) records a number of steps that, according to him, led to a pragmatic account of observation sentences: 1) “The early positivists assumed that observational terms refer to subjective impressions, sensations, and perceptions of some sentient being.” (No references given, but Carnap (1928) is mentioned beforehand.) 2) “Physicalism for some time retained the idea that a scientific

To be sure, Feyerabend is not doing history of philosophy of science (“for our present purpose a brief outline must suffice” (Feyerabend 1962, 35) indeed) and it would be an hollow exercise to measure the correctness of his passing remarks from the vantage point of contemporary historical studies. In a later paper, he would stress again and even more sweepingly the supposed origin of PTO: “The new theory of observability [...] ([...] which was formulated very clearly in the early thirties by Popper, Carnap and Neurath) may be called the *pragmatic theory of observation*)” (Feyerabend 1965b, 243). Yet the wide range of diverging views in the Viennese debate makes it difficult to ascribe a common theory to Carnap, Neurath and Popper.<sup>17</sup> All the same, Feyerabend’s narrative in his (1962) is specific enough to pinpoint some features of the radical development he cherished.

The background to Feyerabend’s narrative is the steady radicalization of Carnap’s and Neurath’s views in their mutual interaction since the publication of Carnap’s *Aufbau* in Vienna (and, starting 1931, between Prague and Vienna), leading to the formulation of syntacticism and physicalism. The tipping point starting this development, as told by Carnap himself,

---

theory should be based upon experiences, and that the ultimate constituents of experience were sensations, impressions, and perceptions.” (Referenced: Carnap (1932d)) 3) “Later, however, a behavioristic account was given of these perceptions to make them accessible to intersubjective testing. Such a theory was held, for some time, by Carnap and Neurath.” (Referenced: Carnap (1932d)) 4) “Soon afterwards the idea that it is experiences to which we must refer when trying to interpret our observation statements was altogether abandoned” and “Popper [...] has been responsible for this decisive turn”. (Referenced: Popper (1959)) 5) “The descriptive terms of Carnap’s ‘thing-language,’ too, no longer refer to experiences.” (References: Carnap (1936), Carnap (1937)) Cf. the similarity and differences to the reconstruction in (Uebel 2007, 442): *Carnap<sub>1</sub>*: the evidential statements concerned, at the bottom, recollections of similarity relations between experiences. *Carnap<sub>2</sub>*: the primitive protocol statements were statements about the phenomenal given variously describable. *Carnap<sub>3</sub>*: the protocol statements were statements about the given that could take various phenomenalistic or non-phenomenalistic forms. *Carnap<sub>4</sub>*: the protocol statements referred to concrete physical states of affairs: psychological states, observed processes or unobserved processes inferred from observed processes. *Carnap<sub>5</sub>*: the protocol statements referred to observable physical states.

<sup>17</sup>In a letter to Kuhn in the early 1960s Feyerabend mentions Carnap only: “By the way, a theory which would allow such procedure has been sketched, in 1932, by Carnap himself and I have discussed it under the title of “the pragmatic theory of experience” in my explanation essay (1962)” (Hoyningen-Huene 1995, 383).

coincides with the opposition of some members of the Vienna Circle to Wittgenstein's view that one cannot talk about language:

In opposition to this view, first tentatively, then more and more clearly, our conception developed that it is possible to talk meaningfully about language and about the relation between a sentence and the fact described. Neurath emphasized from the beginning that language phenomena are events *within* the world, not something that refers to the world from outside. Spoken language consists of sound waves; written language consists of marks of ink on paper. Neurath emphasized these facts in order to reject the view that there is something "higher", something mysterious, "spiritual", in language, a view which was prominent in German philosophy. I agreed with him, but pointed out that only the structural pattern, not the physical properties of the ink marks, were relevant for the function of language. Thus it is possible to construct a theory about language, namely the geometry of the written pattern. This idea led later to the theory which I called "logical syntax" of language. (Schilpp 1963, 29)

Behind Carnap's idea of a "geometry of the written pattern" lies his adoption of the formal viewpoint of the Hilbert school, according to which a logical language is a system of uninterpreted marks rather than meaningful signs. Carnap extended this viewpoint to all knowledge in general. A series of results in mathematical logic aided Carnap's contention that it was possible to speak about the object-language in the meta-language and, then, object-language.<sup>18</sup> This paved the way for Carnap's *Thesis of Metalogic*, already previewed in (Carnap 1932a, 435), that "all philosophical problems which have any meaning belong to syntax" (Carnap [1937] 2014, 280), i.e. all meaningful philosophical sentences are sentences about the form of sentences. Applied to the notion of observation, the Thesis states that "philosophical conceptions about the nature of observation are, properly expressed, opinions about the syntactic analysis of the statements describing observations" (Oberdan

---

<sup>18</sup>See Awodey and Carus (2006) for a detailed reconstruction of this development.

1993, 10).

As we will see, Carnap's distinction between "structural pattern" and the "physical properties of the ink marks" of written language downplays the role that language as a system of physical (spoken, written) events played in his own work. While for Carnap the physical properties of ink marks may not play a role in spelling out the linguistic (logical) function of those ink marks (as they did in Neurath's project of "physicalization of metalinguistic discourse", cf. Uebel (2007), p.164-165), they become relevant where their empirical role is key to clarify epistemological issues.

While Carnap's project of logical syntax is often taken to be coincident with the former (logical) option, the latter (empirical) option is mostly associated with the work of Otto Neurath. In the case of Carnap this holds de facto—Carnap concerned himself mostly with the logical—but not in principle. For one thing, empirical questions hardly vanished from Carnap's writings. While his analysis is often remembered for exposing traditional philosophical questions as meaningless pseudo-questions, it was not its only outcome. Some meaningful questions turned out to have an analytic *and* a synthetic component, which had to be carefully disentangled. The former were to be treated as problems of logical syntax, the latter in terms of physicalization. Traditional problems were thus not simply annulled, but dissolved into *Wissenschaftslogik* plus a *realwissenschaftliches* part. The problem of the empirical foundation of science (the "basis problem") was one of such questions that turned out to have both components. And on various occasions during the protocol sentence debate—without doubt due to Neurath's influence—Carnap not only acknowledged the synthetic component, but also elaborated on it.<sup>19</sup> The causal characterization of observation sentences, which Feyerabend found congenial, is located in this empirical realm. Indeed, I judge it to be Carnap's most perceptive and detailed analysis pushing in this direction.

---

<sup>19</sup>Carus (2007, 166) dates the origin of this development already to a much earlier time: "From some time in 1924, the aspects of the system that had begun under the auspices of phenomenology were transformed either into empirical statements about the workings of the sensory organs or into explicit statements of meta-logic, later called rules of 'logical syntax'."

In the following, I will focus on what Thomas Uebel calls “phase two” of the protocol sentence debate, in particular on substages two and three (cf. Uebel (2007), ch. 7 and 8).<sup>20</sup> Temporally, this timeframe coincides with the publication of key writings by Carnap (and Neurath) in the 1932 protocol sentence debate. I will limit myself mostly to textual evidence from Carnap’s major articles of this phase: Primarily, Carnap’s paper “Die physikalische Sprache als Universalsprache der Wissenschaft” (1932a).<sup>21</sup> As is well known, *Universalsprache* contains the first exposition of physicalism as the thesis that firstly, the language of science, because of its intersubjective nature, coincides with the physical language; secondly, that the physical language is a *universal* language of science; thirdly, that the protocol language is a sub-language of the physical language.

The very general remarks of *Universalsprache* were flanked by more specific and applied investigations, in particular about how physicalism could be carried out in the case of psychology. A decisive push in this direction was surely given by Neurath. As early as 1928, responding in a letter to Neurath’s review of his (1928), Carnap emphatically acknowledged that the actual scientific language is always a “mixture of crystals and dirt” and that one ought to think hard about “what demands are to be made of scientific concepts and statements while the ‘ideal language’ is not available”, announcing that “a few ideas in this direction are already beginning to form ...” (Carnap to Neurath, 7 October 1928, p. 1, RC 029-16-01, cited in Carus (2007), p. 240). Following up on his intention to pursue this line of inquiry, Carnap approached the “‘dirty’ subject of psychology”, a domain that had already attracted his attention (Carus 2007, 240). The published outcome was his “Psy-

---

<sup>20</sup>Uebel (2007) has worked out in detail how the protocol sentence debate came about, dividing the Vienna Circle discussions into three stages and more refined substages. “Proto-physicalism” emerged in phase two, substage one; substage two and three saw the emergence of mature physicalism, first as a theoretical assertion, then as a proposal, in line with Carnap’s looming Principle of Tolerance.

<sup>21</sup>The English translation Carnap (1934b) is bad (which is ironic since the translation effort was aided by Carnap). Based on the German original, I will provide heavily modified translations and give page references to both.

chologie in physikalischer Sprache” (1932c),<sup>22</sup> exercising physicalism in the particular case of then current psychological approaches.<sup>23</sup> The other major paper, “Über Protokollsätze” (1932d),<sup>24</sup> follows up the discussion of the “basis problem” and is well-known in the context of the protocol sentence debate. Lastly, the subsequent “Erwiderung auf die vorstehenden Aufsätze von E. Zilsel und K. Duncker” (1932b) contains a discussion of two critical replies Carnap had received. Firstly, Edgar Zilsel’s reply to “Über Protokollsätze”; and, secondly, a reply by Gestalt psychologist K. Duncker to “Psychologie in physikalischer Sprache”.

I do not claim that it is necessary, or even possible, to discern a chronological development through the writings towards a causal characterization of observation sentences. Rather, I claim that a basic causal characterization is a common denominator of these articles; and each article provides key ingredients towards a fuller pragmatic theory. The following account is then meant as a synchronic snapshot. Similarly, limiting my reconstruction to Carnap’s writings may reduce the evidential weight of my analysis, but I think that by and large Carnap distills in the 1932 articles most of the features that Neurath and Popper contributed towards a pragmatic characterization of observation sentences.<sup>25</sup> In 1932, Popper’s, Carnap’s and Neurath’s views converge in their rejection of the absolutism of “the given”. Alternatively, all three either preferred or demanded sentences referring to medium-sized objects and events. At the same time, their rejection of “the given” correlated with an anti-foundationalist view of protocol sentences accepting their corrigibility and revisability. Specifically, protocols statements are not generally binding, their acceptance being contingent on further conditions and decisions.<sup>26</sup>

---

<sup>22</sup>I will use the translation provided in Carnap (1959).

<sup>23</sup>See Feest (2017) on the psychological research background of Carnap’s work.

<sup>24</sup>I will use the very good translation provided in Carnap (1987).

<sup>25</sup>This does not imply that they agreed on the background motivations for these common views. The conventionalist element was variously embedded in Neurath’s naturalism, Carnap’s syntacticism and Popper’s Neo-kantian framing.

<sup>26</sup>From here their paths variously intersect and diverge: Carnap agreed with Neurath that *sentence-formation* is a legitimate part of the “basis problem” to be explicated in terms of a relational characteriza-



From these assumptions the claim I want to defend is twofold:

- 1) From the vantage point of Carnap's radical blend of syntacticism and physicalism, protocol sentences are not discernible in terms of syntax, and, therefore, there is no logical distinction between protocol sentences and system sentences in principle. This does not mean, however, that protocol sentences cannot be characterized at all. The syntactical indiscernibility view is simply a consequence of a specific way in which the analytic and the synthetic components of the problem of the empirical basis of science were disentangled. In Carnap's analysis, the characterization of protocol sentences is not a logical, but an empirical task: protocol sentences *are* discernible qua physicalization. Incidentally, this seems to me the reason why the syntactical indiscernibility view was not perceived by Carnap as a threatening consequence: an empirical characterization is characterization enough.
- 2) The specific result sketched in (1) is what Feyerabend refers to when he claims that a pragmatic or causal theory of observation can be found in "the writings of the thirties". In particular, Feyerabend's view that "an observation sentence is distinguished from other sentences of a theory not by its content, but by the cause of its production" (Feyerabend 1962, 36) matches the view that protocol sentences cannot be distinguished from system sentences logically, but they can be distinguished empirically.

The next two sections will be devoted to justifying these claims. In doing so, I will mainly

---

tion between an organism's utterances and its surroundings (cf. Carnap's "B-sentence" (Carnap 1934b, 92, revised translation; cf. Carnap 1932a, 460; also Carnap 1959, revised translation, p. 92; cf. Carnap 1932c, 139). Popper, on the other hand, rejected these (empirical) considerations as "psychologistic". On this point Feyerabend clearly takes side with Carnap and Neurath against Popper. But, while Neurath explicitly codified this "sensory condition" in his idiosyncratic sentence acceptance schema, Carnap went with Popper's less regimented requirements for protocol sentences, which were no longer restricted in their syntactic form, which Feyerabend also took to heart. (See Uebel 2007 for an exhaustive reconstruction of these dynamics.)

A more detailed consideration of the specific features and limits of Neurath's and Popper's contributions to a causal theory of observation will be developed in Kuby (forthcoming).

focus on the first one by offering a systematic reconstruction; once the first claim is established, the second is quite apparent. To make a more precise case for the second claim, however, systematic similarities only go so far and I will supplement the discussion with a few remarks concerning the development of Feyerabend's PTO.<sup>27</sup>

### **3. A pragmatic theory of observation in the 1930s protocol sentence debate**

#### **3.1 The detector model of protocol sentences**

At the center of a systematic reconstruction of Carnap's causal theory of observation lies the detector model, which he first introduces as a particular example (Carnap 1932c, 127–8; cf. Carnap 1959, 183–5). Let a body be called a detector (“Detektor”) iff it reacts to a physical event in the surrounding environment with some other physical event. If the correlation is stable enough, i.e. if the first event plays a causal role in bringing about a certain reaction, the latter functions as an indicator (“Anzeichen”) of the former. Carnap introduces the example of a body which increases its electrical conductance when impinged with a certain illuminant. This is a detector of said illuminant. We do not know the body's internal structure (“Innenbau”) and so cannot yet explain its behavior. Still, we can make reliable inference with regard to illuminants (and even, as a proxy, to other processes, e.g. a distinct photochemical reaction triggered by the illuminant, but which is hard to ascertain by itself) by checking the raised conductance level (1932c, 127).

---

<sup>27</sup>Does this informative result vindicate Feyerabend's claim that Carnap's proposal was essentially what his own PTO amounted to? I will give some textual evidence to this end, but in order to judge the degree of similarity between this theory and Feyerabend's PTO one needs a full account of the latter. I suggest that, following a HOPOS-driven approach, the systematic question can be investigated as a historical, genealogical question: ‘Is Feyerabend's PTO a development of Carnap's?’ which will be carried out in Kuby (forthcoming).

Carnap notes that the specific example can be widened into a more general model: “Whether *the detector is organic or inorganic* is irrelevant to the epistemological issue involved.”

He explains:

The function of the detector is basically the same whether we are dealing with a physical detector of specific sorts of illuminant or with a tree-frog as a detector of certain meteorological states of affairs or (if one may believe the newspapers) with a sniffing dog as a detector of certain human diseases. People take a practical interest in meteorological forecasts. Where barometers are not available they may, consequently, use a tree frog for the same purpose. But let us be clear about the fact that this method does not determine the state of the tree frog’s soul [Seelenzustand], but a physically specified weather condition, even if one cannot describe this condition in terms of the concepts of systematized physics. (Carnap 1959, 184)

The ingenious move on Carnap’s side is to subsume the human case under this model, i.e. to argue that the epistemological situation is one and the same in the case of a specific class of organic detectors, namely, human observers: a protocol sentence produced by an observer is nothing but a specific verbal reaction of a (human) body to some event in its surroundings. Indeed, the salient point of a protocol sentence is its role as an indicator of some physical event. To function as an indicator, a correlation between the verbal reaction and a physical event has to be established. If the correlation is stable enough, i.e. if the event plays a causal role in bringing about the verbal reaction, then one can use the verbal reaction as an indicator of that event. In this way, the verbal reactions becomes a “protocol sentence”.

The epistemological situation, whether we are dealing with barometers or with human observers, is the same. Carnap acknowledges a difference in degrees, but not in principle (Carnap 1959, 184; cf. Carnap 1932c, 129).

## 3.2 Physicalization

The first step towards a causal characterization of protocol sentences was the hard won physicalist insight that all there is to a protocol sentence, even the one expressing subjective experiences, is exhausted by its physical interpretation—lest it become inaccessible to intersubjective testing (Carnap 1932a). Here the development of the notion of physicalization (“Physikalisierung”) is key for tracing how the logical-empiricist causal theory of observation (sentences) came about.

In *Universalsprache* Carnap still introduces protocol sentences as “statements belonging to the primitive protocol” (Carnap 1934b, 44), characterized by subjective access to the given and the presence of “qualitative determinations” (Carnap 1934b, 45), yet each and every protocol sentence can also be translated into a sentence of the physical language. A physical language consists of sentences “of the simplest form” (Carnap 1934b, 52) which “express a quantitatively determined property of a definite position at a definite time” (Carnap 1934b, 52–53). Physicalization consists in the correlation of (a class of) physical determinations to certain qualitative determinations. Carnap gives an account of how such a correlation works for different situations (Carnap 1934b, 60–65): a subject *S* can ascertain under which physical determinations she experiences a certain qualitative determination (personal determination); a subject *S*<sub>1</sub> (e.g. a psychologist) can ascertain under which physical determination a subject *S*<sub>i</sub> (e.g. a test subject) experiences a certain qualitative determination (determination by other persons); further, subjects *S*<sub>1</sub>, *S*<sub>2</sub>, ... (e.g. different psychologists) can ascertain under which physical determinations a subject *S*<sub>i</sub> (e.g. a test subject) experiences a certain qualitative determination (determination on other persons made by several experimenters).

“To experience” is an expression in the material mode of speech. Expressed in the formal mode, a subject *S* ascertains which physical determinations are correlated with a protocol sentence of a test subject *S*<sub>i</sub> containing a qualitative determination. (The same holds for the other two situations.) Most interestingly, Carnap immediately addresses the question of how such correlation is possible: “That determinations of this kind are always possi-

ble in principle is due to the fortunate circumstance, an empirical fact, not at all necessary in the logical sense, that [*formal mode*:] the protocol [or, *material mode*:] the content of experience has a certain ordered configuration [Ordnungsbeschaffenheit]. This reveals itself in the fact one can build a physical language in such a fashion that qualitative determinations in protocol language are single-valued functions of the numerical distribution of coefficients of physical states” (Carnap 1934b, 61, revised translation; cf. Carnap 1932a, 445).

Note that Carnap explicitly adduces an empirical happenstance to explain the possibility of such a correlation. This was exploited by Schlick to argue that—contrary to Carnap’s intention—physicalism was not a linguistic thesis, but a plain empirical hypothesis after all (Friedl 2013, 3:267–8). In the paper at hand Carnap tries to anticipate this objection: “It must be noted, however, that these facts, though of empirical nature, are of far wider range than single empirical facts or even specific natural laws. We are concerned here with a much more general structural property of experience [...]” (Carnap 1934b, 65, revised translation; cf. Carnap 1932a, 447).

Carnap’s otherwise clear language gets notably murky in this passage. The most charitable reading here is talk of empirical, contingent conditions of possibility. This, indeed, doesn’t make physicalism *per se* an empirical hypothesis: just as the existence of a spatial plane on which physical marks can exhibit a structural pattern doesn’t make the “geometry of the written pattern” an empirical enterprise, so the contingent existence of a “structuredness of experience”—expressed in the material mode of speech—does not turn the thesis of physicalism into an empirical hypothesis.

Unlike the resulting case of logical syntax, however, the resulting procedure of physicalization *is* empirical: “The procedure consists in that  $S_1$  varies the physical conditions (e.g. the combinations of various frequencies of oscillations) and discovers the conditions to which  $S_i$  reacts with the protocol statement containing the qualitative determination in question” (Carnap 1934b, 62, revised translation; cf. Carnap 1932a, 446).

A simple example might be a scientist who manipulates the electromagnetic frequencies in

the visible spectrum until the test subject utters the protocol sentence: “I see green”. Now we can correlate the qualitative determination “green” to “a definite combination of frequencies of electromagnetic oscillations” (Carnap 1934b, 59), which becomes the corresponding physical determination. Note that the protocol sentence here is modelled as a *reaction* of the test subject to certain events. In this way, the protocol sentence becomes an indicator of another physical event. We can now test a hypothesis that makes a prediction containing the above combination of frequencies of electromagnetic oscillations by checking if the test subject utters the protocol sentence “I see green” (or, say, if the protocol sentence appears in her written protocol) when she is exposed to such physical circumstances.

The main feature of this procedure finds its way into the detector model: the connection between a protocol sentence and what the protocol sentence is an indicator of is now spelled out in terms of simple *causal relations* between physical events. This is the first, decisive step towards a causal characterization of observation.

### 3.3 Protocol sentences as physical events and as linguistic entities

Spelling out the relation between a protocol sentence and what the protocol sentence is an indicator of in terms of causal relations between physical events, means that the protocol sentence itself is, well, a physical event. This *prima facie* modest proposition is indeed a major one, for it does not simply claim that such a sentence is *also* a physical event, but that a protocol sentence exhausts its indicative function qua physical event (an utterance), not qua linguistic entity (say, a proposition).<sup>28</sup> And, as far as the detector model is concerned, the indicative function is all there is to define the epistemological function of uttering protocol sentences.

As Carnap puts it, protocol sentences have to be introduced, at first, not “as scientific sentences, but as scientific facts“. Protocol sentences are utterances, a kind of verbalizing

---

<sup>28</sup>Note that this is not the same as its physicalist interpretation. A physicalist interpretation claims a protocol sentence is about some physical events. The further assertion is that the indicative role is carried out by the protocol sentence as a physical event, not a linguistic entity.

behavior, on par with any other kind of behavior. He expands in painstaking detail on this point while arguing for the fundamental symmetry between introspective reporting and observation of other experimental subjects:

The introspective statements of a psychologist are not, in principle, to be interpreted any differently from the statements of his experimental subjects, which he happens to be reporting on. [...] Further, the statements of an experimental subject are not, in principle, to be interpreted differently than his other voluntary or involuntary movements—though his speech movements may, under favorable circumstances, be regarded as especially informative. Again, the movements of the speech organs and of the other experimental subject’s body parts are not, in principle, to be interpreted differently than the movements of any other animal [...].<sup>29</sup> The movements of an animal are not, again in principle, to be interpreted any differently than those of a voltmeter [...]. Finally, the movements of a voltmeter are not, in principle, to be interpreted differently than the movements of a raindrop [...]. In all these cases, the issue is basically the same: from a specific physical sentence, other sentences are deduced by a causal inference, i.e. with the help of general physical formulas—the so-called natural laws. (Carnap 1959, 195, revised translation; cf. Carnap 1932c, 140–1.)

Referring to the problematic state of psychology as a scientific enterprise, Carnap diagnoses a widespread “confusion of facts in the form of sentences with the sentences themselves

---

<sup>29</sup>Earlier in the article, Carnap notes how animal psychology led psychologists to a physicalist attitude with regard to human utterances: The proponents of behaviorism were led to their position through their work with animal psychology. In this domain, where no statements but only inarticulate behavior can be observed, it is most easy to adopt the right attitude. Sure enough, from this attitude one gets to the correct interpretation of the statements of [human] experimental subjects, conceiving of these statements as acts of “verbal” behavior, which are in principle no different from other behavior. (Carnap 1959, 181, revised translation; cf. Carnap 1932c, 124) .

considered as parts of the system of science” (Carnap 1959, 195). This confusion seems to stem from the widespread belief that the verbalizing behavior of human detectors (introspective statements in particular) works differently than non-human detectors, in that the verbalizing behavior of human detectors is indicative of a qualitatively different notion of experience. This is not the case if we take physicalism seriously:

In using someone’s introspective statement about the state of his own mind (e.g. A’s statement: “A is excited”), the statement, taken as an acoustic event, is the indicator; under favorable conditions, which are frequently satisfied in scientific contexts, the state of affairs referred to is such that it can be described by a sentence (“A is excited”) of the very same form as the acoustic event which functions as an indicator of it. [...] This similarity in form [Formgleichheit] of the acoustic fact and the scientific sentence which is to be inferred from it explains the ready development of the confusion and its obstinate persistence. The muddle which this confusion leads to is cleared up as soon as we realize that here, as in the other cases cited, the issue is only about drawing an inference from an indicator to the state of affairs that is being indicated [angezeigten Sachverhalt]. (Carnap 1959, 195, revised translation; cf. Carnap 1932c, 141)

The difference between “A is excited” qua acoustic event and “A is excited” qua linguistic entity, i.e. a syntactically correct (*syntaxgemäße*) sentence, is paramount to physicalism and is the second step towards a causal theory of observation.

In *Protokollsätze* (1932d), Carnap explores this by introducing an inorganic detector, “a machine which reacts in certain situations by displaying signal-disks. The disks might bear the numerals ‘1’, ‘2’, etc; instead of this they could also be arbitrary meaningless characters.” (Carnap 1987, 458; cf. Carnap 1932d, 216.) Through observation one ascertains that “the two signals ‘1’ and ‘4’ are only jointly visible if it is currently raining lightly outside; ‘1’ and ‘5’ when it is raining hard; [...]” (458; cf. 216) etc. and thus constructs a dictionary



relating the the numerals and meaningless characters, respectively, to one's environment. By doing so, "we can translate certain signal combinations into sentences of our language, e.g., '1,5', into "It is raining hard" (459, cf, 216). Carnap stresses, again, that "whether it concerns a machine which is so built as to give signals, or any other object exhibiting regular observable reactions under specified conditions, makes no fundamental difference. [...]. Even here we can use the reaction as a signal and construct a system of rules for the translation of the signals into sentences of our language. Basically the same situation is present when we find a human who reacts in certain circumstances with specific verbalizations" (Carnap 1987, 459).

Carnap follows through with the separation by introducing "a man whose speech sounds do not belong to any known language" (Carnap 1987, 459). Here we can finally appreciate in full the difference between acoustic event and sentence, for the latter is *not* of the same form as the former. In complete accordance with the first example, we may construct a dictionary by observing the man's utterances in relation to events in his environment: "The man says: 're bim' if it is raining lightly; 're bum' if it is raining hard; [...]" (Carnap 1987, 459), so that "re bim" translates to "it is raining lightly"; "re bum" translates to "it is raining hard", and so on.

The crucial point is that the utterance "re bim" is not (yet) a linguistic statement. It *is treated like* a statement once we can translate it into another language: "With the help of this dictionary we can translate certain sound sequences of the man—we call them *then* statements—into sentences of our language [...]" (Carnap 1987, 459, emphasis added)

Carnap can bring the point home: "The signals of the machine and the statements of the man are treated like sentences of a language in that translation rules are constructed for them. We call them therefore "protocol sentences" of the "protocol language" of the machine or of the foreign man and distinguish this language from the language of our system. Generally every observable process (of a machine, of a man or of whatever), for which a translation rule has been constructed, is valid as a protocol sentence" (Carnap 1987, 459).

The gap between acoustic event and linguistic entity has not been magically bypassed, to

be sure. Not the acoustic event, but its transformation into another language via translation rules determines the content of a protocol sentence. As Carnap has tried to show all along, that “other” language is, ultimately, the physical language.

Most importantly, whether the protocol sentence has meaning is solely determined by syntactical rules, not by (the conditions of) its occurrence. In *Psychologie*, Carnap introduces a striking example to explain the difference:

Should unusual, brilliant patterns suddenly appear in the sky—even if they took the form of letters which seemed to compose a sentence—science could not comprehend them except by first conceiving them, describing them, and explaining them (i.e. subsuming them under general causal sentences) as physical facts. The question whether such an arrangement of symbols constitutes a meaningful sentence must be decided without taking into consideration whether or not it appears in the sky. If this symbol-arrangement is not a meaningful sentence beforehand, it cannot become one no matter how bright an appearance it makes in the sky. Whether a sentence is true or false is determined by empirical contingencies; but whether a sentence is or is not meaningful is determined solely by the syntax of language.

It is no different in the case of those acoustic phenomena that are emitted from the mouths of certain vertebrates. They are first of all facts, physical occurrences, and specifically, sound waves of a certain sort. We can, further, also interpret them as symbols. But whether or not such an arrangement of symbols is meaningful cannot depend on its occurrence as an acoustic phenomenon. (Carnap 1959, 179–80, revised translation; cf. Carnap 1932c, 122)

This is the second step finding its way into the detector model: if in step one the connection between a protocol sentence and what the protocol sentence is an indicator of has been spelled out in terms of *causal relations* between physical events, we now see that the protocol sentence works as indicator qua physical event; any additional function, including a

linguistic function, is independent of the circumstances and specifics of its production.

### 3.4 Indicators of what?

Up to this point mainly the correlation between a protocol sentence and a distal stimulus (the thing sensed) has been considered. As Carnap notes though, there are different ways to fix what the physical event ‘protocol sentence’ is an indicator of<sup>30</sup>. After internal discussions with Neurath, Feigl and Hahn, Carnap already acknowledged that, at the bottom, the choice concerning what a protocol is an indicator of is conventional:

We can distinguish at least five events, Carnap wrote after this meeting: (a) the thing sensed, e.g. a tree; (b) the retinal impression of the tree; (c) the corresponding trace of that impression in the visual processing part of the brain; (d) the corresponding intent, in the linguistic centres of the brain, to impart a report of that impression; and (e) the utterance corresponding to that intent. Whether we regard (e) as an indicator only for (a), skipping the intervening steps, or also of all or any of those (and further) intervening steps also, is a matter of convention, he writes, and will vary according to our purposes. And we can observe empirically that there is a correspondence between these five steps as viewed from a subjective point of view and the same steps viewed as a physical causal chain. (Carus 2007, 244, detailing Carnap’s “Zur phänomenalen Sprache”, RC 029-17-02)<sup>31</sup>

As usual, “conventional” doesn’t mean “arbitrary”, but “relative to a specific purpose”. Two options were mainly considered in physicalism:

---

<sup>30</sup> As an aside in *Universalsprache*, explicitly in *Psychologie* and *Protokollsätze*.

<sup>31</sup> Carus (2007, 244) further observes that “Carnap was clearly on the way to the pragmatic view of the observation language that later found its context within the new pluralism of ‘On Protocol Sentences’”. The link to pluralism (language-relativity) gives reason to suspect that “pragmatic” may be used in the same misleading way we already encountered in Oberdan’s paper.

1.  $(e) \rightarrow (a)$ : The utterance of a protocol sentence is an indicator of the distal stimulus.
2.  $(e) \rightarrow (c)$ : The utterance of a protocol sentence is an indicator of a particular neuro-physiological state of the brain.

In *Protokollsprache*, these two options are distinguished as T-sentences and B-sentences, respectively, and their relative merits are discussed. Translation rules yielding T-sentences “lead from protocol sentences to sentences which refer to things in the environment of the man [i.e. human detector] in question” (Carnap 1987, 460). They are called T-rules. For example, we may establish that an utterance “re” translates to the T-sentence “It is raining”. B-sentences, on the other hand, “refer to the momentary state of the body B of the person in question” (Carnap 1987, 460). Here the utterance “re” is translated into the B-sentence “The body B is rain-observing”. The B-translation given by Carnap is surprising at first, for we just stated that the second option identifies the indicated event with a neurophysiological brain state! To better understand this issue, we have to take a little detour and go back to an important section of *Universalsprache*, in which Carnap addresses the converse situation: not bottom-up (how protocol sentences become part of the physical language), but top-down: how protocol sentences can be derived from system sentences.<sup>32</sup>

Carnap notes that the “simplest form of such deduction” of a protocol sentence from physical sentences obtains if the latter describe the bodily state of the subject (Carnap 1934b, 85). It is true that from a physiological point of view the bodily state is best defined in terms of “the state of the central nervous system and especially the brain”, but—Carnap admits—we don’t know enough about to define the physical determination in such a way. But this is not a show-stopper for we can already read external bodily cues, even if in a

---

<sup>32</sup>A fact he establishes beforehand; see (Carnap 1932a, 440): having explained that the “sentences of the system of science [...] are not, in the proper sense of the word, derived from protocol sentences” for “from no collection of protocol sentences, however many, can [a singular statement] be deduced, but it is in the most favourable case continually supported by them”, Carnap notices “the deduction is possible but *in the converse direction*” (Carnap 1934b, 48–49, emphasis added): protocol sentences can be deduced from system sentences, which makes it possible to test the theory by checking if the sentence appears in the protocol or not.

more primitive way. Let the physical sentence P be “The body S is now red-seeing”. The bodily state “red-seeing” may correspond to a specific neurophysiological configuration, which would give the physical determination to physicalize the qualitative determination “red” occurring in the protocol sentence p “(S sees) red now”. Not having enough physiological knowledge to correlate the qualitative determination with a neurophysiological state of S, we have nonetheless other bodily cues which can serve as physical determination, namely, the class of all bodily stimulus-responses that we usually take as indicators to tell that someone “sees red now”:

Let us, e.g. denote by ‘seeing red’ the particular state of the human body characterized by the fact that certain specified (physical) reactions appear in response to certain specified (physical) stimuli. (For example, stimulus: word sound “What do you see now?”, reaction: speech movement “red”; stimulus: word sound “Point to the color you have just seen on this chromaticity diagram”, reaction: the finger points to some definite spot on the diagram; etc. Here we should list all those reactions that we usually regard as distinguishing marks to tell that someone is “seeing red now”). It is still true that we do not know the numerical distribution of the physical coefficients which characterize the human body in this state of “seeing red” but *we do know many physical events which often occur either as cause* (e.g. bringing a poppy before the eyes of the person concerned) *or as effect of such a state* (e.g. certain speech movements; halting movements under certain conditions.) Hence we can, for one, recognize that a human body is in that state; secondly, we can gain from it predictions about further bodily states that can be expected. (Carnap 1934b, 86–87, revised translation, emphasis mine)

This strategy is central for Carnap to counter the objection from the “undeveloped state of physiological knowledge” (*Unbekanntheit der Physiologie*): “Our current knowledge of physiology—especially our knowledge of the physiology of the central nervous system—is not yet sufficiently advanced to enable us to know to what class of physical conditions

something like excitement corresponds. Consequently, when today we use the sentence ‘A is excited,’ we cannot mean by it the corresponding physical state of affairs” (Carnap 1959, 175, revised translation).

While Carnap concedes that we can not translate “the sentence “A is excited” into a sentence of the form ‘such and such a physico-chemical process is now taking place in A’s body’ (expressed by a specification of physical state variables and by chemical formulae)” just yet, it does not follow that the sentence does not correspond to a physical state: the sentence “A is excited” indeed “refer[s] to the physical structure of A’s body—though this structure can only be characterized by potential perceptions, impressions, dispositions to react in a specific manner, etc., and not by any specification of state variables. Our ignorance of physiology can therefore affect only the mode of our characterization of the physical state of affairs in question. It in no way touches upon the principal point: that sentence [“A is excited”] refers to a physical state of affairs” (Carnap 1959, 174–5).

This point is repeated time and again, and it applies to inorganic and organic detectors alike.<sup>33</sup>

We can now come back to understand the issue about *direct* and *indirect* B-translation. In the ideal case, B-sentences may describe the neurophysiological state of B. However, the lack of physiological knowledge prevents us, for now, from making this determination.

---

<sup>33</sup>In the case of the illuminant detector: This practical application will not be impeded by our ignorance of the detector’s micro-structure and our inability to explain its reaction in physical terms. In spite of our ignorance, we can certainly say that the detector isolates a certain physically specified class of illuminants. The objection that this is not a physical class since we cannot characterize it by a specification of optical state coordinates but only by the behavior of the detector will not stand. (Carnap 1959, 183) Like the case of our ignorance of the neurophysiological configuration, in which the class of bodily reactions is already a *physical* specification, here the conductance level increase specifies the physical characterization. The point is made again in the case of sentences containing the psychological concept “excited”, which are content-equivalent to a physical sentence about a bodily state, “whereby the physical term:”physically-excited” is defined as a disposition to react under certain circumstances in a specified way: “Under such and such circumstances, such and such gestures, expressions, actions, and words occur” (Carnap 1959, 193).

Like the case of “red-seeing”, the bodily state “rain-observing” corresponds to a specific neurophysiological configuration, which would directly give the physical determination of the bodily state. Not having enough physiological knowledge to correlate “rain-observing” with a neurophysiological state of B, we can however investigate the physical events which occur either as cause (in the environment of B) and as effect (the behavioral dispositions of B) of such a state. This yields a B-translation that exhaustively characterizes the bodily state of B in terms of its causal history and its observable behavioral effects.

To be sure, Carnap doesn’t see this behavioral stimulus-reaction characterization as the end goal. It is a makeshift. Though “there is no difference of kind between the two characterizations, [...] the difference of degree is indeed sufficiently great to give us a motive for pursuing the empirical investigations which might bring the direct physical characterization within our grasp” (Carnap 1959, 184). In other words, it is a way to deal with the incomplete state of current science—looking towards a less incomplete one.<sup>34</sup> On the other hand, with this investigation Carnap is acknowledging Neurath’s admonition “that what we have before us is always a mixture of crystals and dirt” and that we have to make do “while the ‘ideal language’ is not available” (Carnap to Neurath, 7 October 1928, p. 1, RC 029-16-01, cited in Carus (2007), p. 240). For actual use cases, if the B-translation is used, it will be used in this *indirect* manner.<sup>35</sup>

The B-translation is the third step towards a causal characterization of observation sentences. After having introduced the physical event ‘protocol sentence’ and shown its indicative function, we now possess an account of what the protocol sentence is an indicator of.

---

<sup>34</sup>As noted by Crawford (2014), the same point was later made again by Feigl (1958) (reprinted in Feigl (1967)) to the effect that “the progress of the science of psychology will unfold in two stages: an initial “peripheralist” black-box behaviouristic stage and a later “centralist” neurophysiological stage” (Crawford 2014, 716).

<sup>35</sup>But note that Neurath agreed with Carnap on the direction of this development in the internal discussion: “([Agreeing] with Carnap): We can already build a macroscopic, purely behavioristic theory; later its construction may become much easier by the addition of brain processes” (Neurath, “Besprechung über Physikalismus am 4. März 1931”, RC 029-17-03).

### 3.5 Indicative function and reliability

Discussing the merits of the T- and B-translation in *Protokollsätze*, Carnap explains that the difference between the two types of translation rules lies in the degree of reliability of the physical T- and B-sentence thus obtained: “The T-translation is the ordinary one and for practical life the more suitable, since customarily what matters for us is to learn something about the environment of B. The B-translation is the more reliable; it is preferred when we are concerned mainly with certainty, for example, with critical testing. In both cases we use the statement of our neighbor’s B to enrich our knowledge about the processes (physical, intersubjectively comprehensible processes) just as we evaluate the statements of the signal machine for the same purpose” (Carnap 1987, 461).

Why is the T-translation less reliable than the B-translation? Because “observation teaches” that the T-translation “occasionally yields a false sentence (namely in the cases one usually characterizes as dream-statements, hallucinatory statements, lies or the like; in our account, however, we are not yet in possession of these concepts)” (Carnap 1987, 460). It might well happen that our fellow detector utters “re” when it is not raining, so that the T-translation “It is raining” is false. On the other hand, according to the B-translation the utterance “re” does not indicate a distal stimulus, but a bodily state of the detector: “The body S is now rain-observing”. Ideally, we want to be able to characterize “rain-observing” in terms of a specific neurophysiological configuration that only obtains e.g. if a *proximal* stimulus (cf. step (b) in the causal chain) is in fact present—thus excluding certain types of hallucinatory statements. Since, as Carnap already explained, the ideal case is hard to come by, we can increase the indicator’s reliability by letting a detailed causal history and behavioral description enter the characterization of the bodily state. Thus “rain-observing [...] can perhaps be characterized as being found in certain conditions (namely when it is raining) or if rainlike audible or visible processes are present, and the eyes or ears of B are in the appropriate relative position to these processes) and as stimulating such and such observable bodily reactions (e.g., on the appropriate stimulus, perhaps questions, under appropriate circumstances the statement”re“)” (Carnap 1987, 460).



We are left with a curious situation: the indicative function of a protocol sentence seems to be unrelated to the indicator's reliability and vice versa. This is, indeed, a major feature of the causal theory of observation for it divorces the indicative function of protocol sentences from any "intrinsic", non-physicalist connection between the protocol sentence and the thing sensed, often associated in metaphysical views with notions of *certainty* of subjective experience expressed by the protocol sentence.<sup>36</sup>

This is the fourth essential step towards a causal characterization of protocol sentences, complementing the sharp division between the indicative function and the content of a protocol sentence presented in §3.3.

### 3.6 Protocol sentences as observation sentences

The observational function of a protocol sentence uttered by S is achieved when we learn something about the environment of S and not just about the bodily state of S.<sup>37</sup> If we want to learn something about the environment of S, a B-translation seems to miss the mark. Our goal is to infer a T-sentence from the protocol statement of S. As we have discussed at length, a protocol sentence p "A red ball is lying on the table here" may be translated to a T-sentence, e.g. the physical sentence P<sub>1</sub> "On the table lies a red ball (i.e. a ball having such and such physical properties)" (Carnap 1934b, 92, revised translation; cf. Carnap 1932a, 460). In terms of practical usage, the T-translation is potentially unreliable because e.g. the ball could be hallucinated while there's no ball on the table. The unreliability of a T-translation shows that the indicative function is a necessary, but not a sufficient condition to actually learn something about the environment of S (i.e. we might be actually learning something about the hallucinatory states of S).

How can we achieve a reliable observation? Carnap shows a way out by offering a more

---

<sup>36</sup>More precisely, it shows that any experience of certainty accompanying the indication process is unrelated to its reliability.

<sup>37</sup>"Observation" should not be restricted to the visual sense modality. One of the outcomes of physicalism is: all that can be intersubjectively communicated is also translatable in other sense modalities.

refined analysis of the problem. Until now the unreliability has been characterized in practical terms. In *Universalsprache* he spells it out in more precise syntactical terms: the salient point is that  $p$  and  $P_1$  are *not* “content-equivalent” (*gehaltgleich*) (Carnap 1934b, 92; cf. Carnap 1932a, 460), i.e. “mutually deducible.”<sup>38</sup> There is no way in which we can transform  $p$  to  $P_1$  via syntactical rules. This is a pivotal result. It mirrors the result of the analysis in §3.4, i.e. that physical events are related to protocol sentences only empirically (as causal relations), and not logically. Any relation between them is of a contingent, empirical nature.

The B-translation, instead, yields a B-sentence that is content-equivalent with  $p$ , namely, the physical sentence  $P_2$  “The body of  $S$  is now in a physical state  $Z$ , where  $Z$  is defined by a class of conditions such as: following a stimulus “What do you see”, the speech movement “A red ball on the table” obtains as a reaction; and: after a red ball has been put on the table and  $S$  is in a suitable situation,  $Z$  occurs” (Carnap 1934b, 92, revised translation; cf. Carnap 1932a, 460). We see that the physical interpretation of a protocol sentence refers to a bodily state  $Z$  that is exhaustively characterized by a conditional clause of the form “If such and such, then  $p$ ” (Carnap 1959, revised translation, p. 92; cf. Carnap 1932c, 139).<sup>39</sup> If an inference from  $P_2$  to  $P_1$  ought to succeed, the physical event “expressed” by  $P_1$  has to be among the empirical circumstances listed in the antecedent of the conditional clause. (The physical event  $P_1$  is among the empirical circumstances causing the production of  $p$ , just as the utterance of  $p$  is one among many possible physical events caused by the physical event

---

<sup>38</sup>Cf. Carnap (1934b), p. 91: “[T]he content of a sentence is constituted by the possibility of deducing further sentences from it; if for any two sentences one can deduce the same further sentences, the two sentence under consideration have the same content [...], as “[t]he metalogical term “content-equivalent” is defined as “mutually deducible” (Carnap (1934b, 87, revised translation); cf. Carnap (1932a, 458). See also (Carnap 1934a, 38)).

<sup>39</sup>In *Universalsprache* Carnap clearly states that physical sentences in the form of a B-sentence allows the “simplest form” of deduction of a protocol sentence and most importantly, that “all other deductions are more complicated and can be reduced to this case” (Carnap 1934b, 85). Yet in *Protokollsätze* he portrays the T-translation and B-translation as equivalent, implying that there’s no ranking in terms of reducing one to the other.

$P_1$ .) Only then, as Carnap says, “ $P_1$  can be inferred from  $P_2$  under appropriate conditions, by using  $Z$  and suitable natural laws” (Carnap 1934b, 92, revised translation; Carnap 1932a, 460). The inference of  $P_1$  from  $P_2$  is “an inference from an effect to an habitual cause, as is common in physics and in everyday life” (Carnap 1934b, 92, revised translation; Carnap 1932a, 460). In other words, a reliable indicative function of  $p$  is only accessible through a statement about the bodily state  $Z$  of the test subject  $S$  uttering  $p$  when connected to a causal account of the empirical circumstances leading to the production of  $p$ . And the indicator is only as reliable as the causal account defining  $Z$  is, which is an empirical matter.

We see: as far  $P_1$  is concerned, the B-translation offers no shortcut to remedy the T-translation’s unreliability. But it offers a better analysis of the empirical nature of reliability and, thus, it gives an indication of what is involved in actually learning something about the environment of  $S$  through a protocol sentence uttered by  $S$ .

Carnap concludes: “The usual interpretation of a protocol sentence as referring to a certain state of the subject’s surroundings is really an indirect interpretation, consisting of the actual interpretation (as referring to the bodily state) and a causal inference [relating the bodily state to the subject’s surroundings] (Carnap 1934b, 93, revised translation; Carnap 1932a, 460)<sup>40</sup>.

This concise statement summarizes the pragmatic characterization of observation on Carnap’s side.

---

<sup>40</sup>Uebel (2007, 253) argues that “what physical things [Carnap and Neurath] took the protocols to be speaking of” were different: “for Carnap, it was states of the bodies of protocollists, for Neurath, the states of bodies of protocollists in their relation to their environment.” In light of the characterization above, I submit that Carnap was here more “Neurathian” than Uebel portrays. While Uebel’s description of Carnap’s position is correct if the direct B-translation is taken at face value, it is not with respect to the indirect B-translation. From the vantage point of the latter, a protocol is indeed about “the state of the bodies of protocollists in their relation to their environment”, as Uebel puts it, because in the indirect B-translation the characterization of the bodily state of the protocollists is defined by the conditional clause including a detailed account of exactly such a relation.

### 3.7 A pragmatic theory of observation sentences as a solution to the “basis problem”

In his subsequent article *Erwiderung auf die vorstehenden Aufsätze von E. Zilsel und K. Duncker* (1932b), Carnap spells out the radical consequences of a causal characterization of observation sentences with regard to the “basis problem”. Answering Zilsel’s query on how we can go about discerning the “‘true’ protocol sentences” and related system(s) of science, Carnap agrees that “no such discernment is possible by purely logical means [...]” and that “the ‘real system’ (including protocol sentences) has no formal distinguished status” (Carnap 1932b, 179, my translation, emphasis added). We may, however, reinterpret Zilsel’s question—Carnap continues—as an empirical (historical, sociological) question about which set of protocol sentences and which scientific system is actually in use among scientists at a particular time and place. This way, a distinction *is* possible. Indeed, “there is no distinction for ‘our’ scientific system apart from a historical distinction establishing that the latter is the scientific system of our cultural background [Kulturkreis]” (Carnap 1932b, 180, my translation).

Alternatively, and more appropriately in my view, Carnap interprets Zilsel’s query as a “methodological question: ‘Which rule [Vorschrift] should one go by to formulate a ‘correct’ protocol sentence?’” (Carnap 1932b, 181, my translation). Carnap’s laconic answer: “There is no such rule, because such a rule would be trivial, e.g. ‘If you have toothaches, say ‘toothaches’ ” (Carnap 1932b, 182, my translation). An “original” language-learning mechanism (not via translation to another language) is not set up via rules, but imparted via “practical measures”:

Just as one can teach a child with examples and handholding how to spit out cherry pits, so one can habituate him (as a “conditioned reaction”, as behavioral psychologists call it) to say “this house is tall” or “I have toothaches”. If the pupil verbally reacts like us (i.e. he says “the house is red” when the house is red), we call him “correctly speaking”. The protocol sentences of

those “speaking correctly” will be used by the scientific community also in those cases where we can’t (like the house example) check the correctness immediately. *The formation of protocol sentences doesn’t involve pure, but descriptive semantics; in the usual classification: not logic, but zoology, i.e. physics.* Here there’s talk not only of sentence structures, but of those sentences themselves as physical shapes [Gebilde]. From the viewpoint of logic one can only connect sentences (more precisely: sentence structures) with other sentences; from the viewpoint of physics, on the other hand, we can describe how sentences as speech movements are formed under certain conditions, or we can advance stipulations on their formation. (Carnap 1932b, 182–3, my translation, emphasis added).

In a radical move the “basis problem” is reduced to a problem regarding the formation of protocol sentences and, from there, to an issue of behavioral training. This is unapologetically recognized by Carnap: “The whole system of science rests in the end on sentences (namely, protocol sentences) which are not formed via logical operations, but through a specific practical procedure, through trained reactions” (Carnap 1932b, 183, my translation).

Systematically speaking, this empirical vantage point allows us to discern at least two ways to put the detector model to use: we can use detectors by either describing existing reactions; or, if the reactions are not hardwired, we can change existing reactions. In the case of the “machine which reacts in certain situations by displaying signal-disks” (Carnap 1987, 458; cf. Carnap 1932d, 216.), we can describe the dials in relation to the environment and construct a translation such that ‘1’ and ‘4’ indicates “It is raining”; or, if we have access to the dials, we can change them by directly displaying the words ‘It is raining’ in the appropriate situation. The same holds for the “man whose speech sounds do not belong to any known language” (Carnap 1987, 459). We ascertain that he utters “re bim” if it is raining lightly; we may change his conditioned reaction to say “It is raining lightly” instead of “re bim” in the same situation.

Let’s extend this reasoning to the scientific case: a simple voltmeter moves a pointer across

a scale. We can change and tune the pointer to move in proportion to the voltage and add the word “volt” next to each point of the scale. We can change the voltmeter from an analog to a digital interface displaying the words “There is a potential difference of ... volts in the electrical circuit now”. We may even add a speech synthesizer to utter those words.

The same holds for an organic detector: we may find a human detector consistently saying “My tongue is tingling” when pressing a battery against the tip of her tongue. We may then change her verbal behavior to utter the words “There is an electric potential difference” in the same situation. If the tingling increases in proportion to the voltage, she may be trained to utter words containing qualitative (“There is a low potential difference”) or quantitative (“The potential difference is 5 volt”) indications. This situation can be complicated at will to describe actually occurring observation processes in science: a scientist uttering the words “5 volts” when the pointer of the voltmeter shows “5 volts” on the scale is just a further conditioned reaction and a further physical event in the causal chain. Adding even more complexity, a scientist may also be her own detector (cf. Carnap (1959), p. 184.) Sure enough, whether we want to interpret the scientist’s utterance as an indicator that such an indication appeared on the voltmeter or as an indicator that there was a potential difference of 5 volts in the electric circuit is a matter of convention, depending e.g. on whether we are calibrating the voltmeter or making a reading of the circuit. Indeed, most scientific observations involve multiple causally-linked detectors—last, but not least, the scientist herself.

The last example shows a third way to approach the detector model: if the detector is susceptible to new, variable inputs (if it can detect new physical events in the surrounding environment), one can train new reactions (what Carnap calls “original” (*ursprüngliches*) language-learning). Most of the conditioned reactions relevant to the scientific community are not part of the generally available bodily dispositions of human observers. Scientific observation is chiefly dependent on the training of new and reliable reactions, imparted via practical measures, to new physical events.

## 4. Feyerabend's PTO in light of Carnap's detector model

Feyerabend's own contribution to a pragmatic theory of observation was first developed in his dissertation (1951), in which he proposed a forerunner of PTO he called a "theory of reactors" (*Theorie der Reaktoren*) (Feyerabend 1951, 60, my translation). Here he would begin where Carnap left off: to "study how sentences as speech movements are formed under certain conditions" and to "advance stipulations on their formation", as Carnap (1932b, 180) envisaged. (In his paper (1958), Feyerabend would first publish a short—and much abridged—summary of the stipulative task only.<sup>41</sup>)

In the dissertation, the *Theorie der Reaktoren* is explicitly presented as an "empirical theory, which resembles in many ways a theory of thermometers" (Feyerabend 1951, 60, my translation), in complete agreement with Carnap's account. Feyerabend's "reactor" is evidently a synonym for Carnap's "detector". Feyerabend's theory models observers as reactors (viz. detectors), measuring devices which react in a certain way under certain conditions. Already in this early model, the utterance (more generally: the production) of protocol sentences is one of the possible observer's behavioral patterns, which are *causally* determined by the observer's surrounding environment.

Feyerabend's model gives an empirical theory for the basic empiricist tenet that an (appropriately calibrated) detector tells us something not just about its own state, but also about what is detected in its surroundings. What it tells us, however, is rather minimal: that something exists. That is the most the detector model can achieve. It can tell us that *something exists*, but *not what exists*. But this limitation is not a disadvantage. Indeed, one of the features of PTO, which we have detailed in §3.3 and that Feyerabend took to heart, is that it sharply distinguishes between the indicative function (i.e. that something exists) and the linguistic function of an observation sentence—its interpretation (i.e. what

---

<sup>41</sup>This part would eventually reappear briefly as the conception of "quickly decidable sentences" in Maxwell (1962), p. 13. But note that Feyerabend already presented the pragmatic characterization in print beforehand, cf. Feyerabend (1954), p. 471, recently translated as Feyerabend ([1954] 2015).

exists), and that the latter cannot be derived from the former. Later, Feyerabend (1958, 145) tracks this sharp separation by distinguishing between “sentences” (Carnap’s “acoustic events”) and “statements” (Carnap’s “syntactically correct sentence”). Here the terms “observation-sentence” and “observation-statement” are used to specify the respective conditions for the uninterpreted and the interpreted status of a detector’s indication: “[T]he pragmatic conditions concern the relation between observation-*sentences* (*not* statements) and human beings without making any stipulation as to what those sentences are supposed to assert. Further conditions will have to be added if we want to obtain a fully-fledged language. Any complete class of such further conditions will be called an *interpretation*. A particular observation-language is completely specified by its characteristic [i.e. the set of pragmatic conditions] together with its interpretation. The distinction between the pragmatic properties of a language and its interpretation is clear and unambiguous” (Feyerabend 1958, 145–6).

The ‘epistemological’ lesson that Feyerabend’s PTO inherits from Carnap’s account of observation sentences is that while the stipulations of the logical syntax make protocol sentences meaningful, they cannot discern whether a particular sentence is a protocol sentence. Only a detailed empirical account of the causal chain—viz. the conditioning procedure (resulting in a list of “pragmatic conditions”)—can establish the causal role of a physical event in the production of a verbal reaction, in which case we call it a “protocol sentence”. (In case of protocols sentences indicating some physical event in detector’s surroundings, we may call them “observation sentences”.) Or, in Feyerabend’s words, which we are now able to fully understand: “an observation sentence is distinguished from other sentences of a theory not by its content, but by the cause of its production, by the fact that its production conforms to certain behavioral patterns” (Feyerabend 1962, 36). Conversely, no conceptual analysis of the detector’s verbal reaction, or even of the detector qua physical system, however detailed that analysis may be, will give us any indication as to whether the observation sentences are meaningful or not, nor what their meaning is.

It may be appropriate here to dwell on the last point a little longer: Like Carnap’s detector



model, PTO leaves completely open the question of how we can establish what the meaning of observation sentences is. Feyerabend would eventually propose an answer by introducing his “thesis I”, according to which “*the interpretation of an observation language is determined by the theories that we use to explain what we observe, and it changes as soon as those theories change*” (Feyerabend 1958, 163).

But this proposal, thesis I, is independent from PTO. Contrary to Oberdan’s contention that thesis I is the “first idea” of PTO and that the causal account follows from it (cf. §2.1), the question of meaning is not part of (or answered by) a pragmatic theory of observation. A case in point: Carnap’s answer, a syntactical deflation of meaning, differs from Feyerabend’s. It is notoriously difficult to give a precise account of the scope of Carnap’s syntax, because his notion of syntax includes elements of—what he would later call—semantics. In any case, “interpretation” is not a syntactical term in 1932. Here the very translation of a protocol sentence into a physicalist language would make the protocol meaningful—and the meaning would be given by the syntactical rules of the language. In Feyerabend’s case, thesis I is sometimes identified with his “contextual theory of meaning”: “the meaning of every term depends upon the theoretical context in which it occurs. Words do not ‘mean’ something in isolation: they obtain their meanings by being part of a theoretical system” (Feyerabend 1965a, 180).<sup>42</sup>

But whether this identification stands up to scrutiny is a matter of debate.<sup>43</sup>

To be sure, Feyerabend is not innocent in muddying the waters, as he presents time and again PTO and thesis I in conjunction. But this should be interpreted, I submit, as the notion that PTO enables thesis I in two ways: it leaves open the question of how the interpretation of observation sentences arises, so that thesis I can be introduced at all; and it gives a negative answer to competing solutions: the “pragmatic” principle of meaning, which would locate the interpretation at the pragmatic level PTO is concerned with, identifying the meaning of an observation sentence with its linguistic use; and the “phenomenological” principle

---

<sup>42</sup>But see already Feyerabend (1962), p. 88.

<sup>43</sup>See Oberheim (2006, 58–63) for a thorough discussion of the relation between the two.

of meaning, according to which meaning is determined by what “the given”, the sort of introspective capacity in the content of experience that the physicalist turn had rejected.<sup>44</sup> And, indeed, we observe in Feyerabend’s writings that the choice of thesis I over alternative accounts of meaning is not uniquely determined by assuming PTO. This choice is argued by Feyerabend affirmatively by appeal to epistemic and even ethical norms, which—according to him—should compel us to assume thesis I over alternative accounts.

Finally, Feyerabend not only developed further Carnap’s detector model, but also applauded Carnap’s contribution as a decisive step in the development of empiricism. As he later remarks:

In the case of measuring instruments, the pragmatic theory degenerates into a triviality: nobody would ever dream of asserting that the way in which we interpret the movement of, say, the hand of a voltmeter is uniquely determined either by the character of this movement itself or by the processes inside the instrument; [...] Taken by themselves the indication of instruments do not mean anything unless we possess a theory which teaches us what situations we are to expect in the world, and which guarantees that there exists a reliable correlation between the indications of the instrument and such a particular situation.  
(Feyerabend 1962, 36–37)

It is its extension to human observers that is novel—and bold. This is, I argue, the main advance that Feyerabend looked up to: Carnap dispenses with what Feyerabend calls the “assumption that human observers are something special and cannot be treated in the same manner as physical measuring instruments” (Feyerabend 1962, 38) that had survived in empiricism for so long and that the pragmatic theory of observation finally dealt with: “It is not easy to set down in a few lines the reasons for this exceptional treatment of human

---

<sup>44</sup>This terminology was explicitly introduced in Feyerabend (1958, 149). But the notion was already central in Feyerabend’s unpublished dissertation (1951), where it was discussed under the heading “subjective theory of basic statements”.

observers. Nor is it possible to criticize them thoroughly and thereby fully pave the way for the acceptance of the pragmatic theory of observation. However, such a comprehensive criticism is not really necessary here. It was partly given by those very same philosophers who are responsible for the formulation of the pragmatic theory (Footnote: Cf. Carnap 1932a; Carnap 1932c) (which most of them dropped later on, their own excellent arguments in favor of it notwithstanding)” (Feyerabend 1962, 37–38).

This resonates quite nicely with Carnap’s remarks about the “the emotional resistance to the thesis of physicalism” while discussing the many objections advancing other (non-physical) interpretations of protocol sentences:

Such resistance is always exerted against any thesis when an Idol is being de-throned by it, when we are asked to discard an idea with which dignity and grandeur are associated. As a result of Copernicus’ work, man lost the distinction of a central position in the universe; as a result of Darwin’s, he was deprived of the dignity of a special supra-animal existence; as a result of Marx’s, the factors by means of which history can be causally explained were degraded from the realm of ideas to that of material events; as a result of Nietzsche’s, the origins of morals were stripped of their halo; as a result of Freud’s, the factors by means of which the ideas and actions of men can be causally explained were located in the darkest depths, in man’s nether regions. (Carnap 1959, 167; cf. Carnap 1932c, 110)

## **5. Concluding remarks on “disinfected empiricism”**

In this article, I reconstructed Carnap’s detector model of observation sentences as an instance of a pragmatic theory of observation, taking my cue from an incidental remark by Feyerabend to the effect that such a theory had been developed by leading Logical Empiricists during the height of the protocol sentence debate. I first proceeded to dispel a misunderstanding concerning the term “pragmatic” raised by Oberdan. As used by Feyerabend,

this adjective does not refer to the conventional nature of the noun it modifies. Instead, it refers to a specific semiotic relation between users of a language and their environment describable by empirical means (specifically, the causal relation between verbal behavior of language users and their environment). This usage of the term “pragmatic” is also in line with Morris’ and Carnap’s use. Secondly, I agreed with Oberdan’s claim that no pragmatic theory of observation (in the misunderstood sense) can be found in Carnap. But, building on the intended meaning of “pragmatic”, I reconstructed how a pragmatic theory of observation sentences can indeed be found in Carnap’s writings. In particular, I sketched such theory by reconstructing it in terms of a detector model that interprets observation sentences as bodily dispositions indicating some physical events in the surroundings of the detector. Finally, I showed how Feyerabend’s PTO picks up central features of Carnap’s account and also shares some of the same motivations.

A few concluding remarks on empiricism are in order. Feyerabend gives a specific reading of Carnap’s account as a crucial step to a “disinfected empiricism”; in fact, “the empiricism of the thirties was disinfected in the sense desired here” (Feyerabend 1962, 31). I take Feyerabend to give a specific interpretation of this move, in which the notion of “experience” relevant to knowledge is divested from any subjectivist interpretation or subjective grounding. Historically, this new conception had resulted from the radical physicalist understanding of experience that Neurath initiated and Carnap accepted without reservation.<sup>45</sup>

---

<sup>45</sup>I do not think that Carnap’s favorable remarks on “methodological solipsism” at his physicalist peak offers a refutation of this point. Instead, our reading gives an interpretation of his remarks that make Neurath’s concerned objections on this point groundless. In short: Carnap’s remarks do not entail a privileged epistemic vantage point of first-person accounts. More than a proposal, his “methodological solipsism” is an observation of a contingent fact: detectors play different roles in the causal chain. The obvious analogy would be the difference between human observers and thermometers. Both are detectors, but it is humans who deliver an interpretation of the thermometer’s indication, not the other way around. But this is based on the contingent fact that inorganic detectors do not provide their own syntactical rules. If thermometers were able to construct their own syntactical rules, the specific difference with human observers would cease to exist. In the same way, the processing of the output of other detectors hinges, contingently, on one’s own ability to be part of the causal chain and these outputs become inputs only if they are processed by one’s own self, which

It was certainly an expression of an “epistemology without a knowing subject”. In particular, no experiential faculties of the human observer are the justificatory endpoint for the “basis problem”. But, crucially, and in contrast to Popper, Feyerabend reminds us that it was after all an *epistemology with a ‘knowing object’*, as it were: from the fact that human observers and their faculties cannot serve as justificatory endpoints, it does not follow that they ought to be thrown out of the equation, as is the case in Popper’s idealistic flight into World Three. A (generalized) anti-psychologistic attitude exhorts us not to conflate logical problems with empirical ones, but this—in Carnap’s case—goes both ways.<sup>46</sup> From the fact that the “basis problem” can be stated in clear terms and that there is no logical grounding to its solution (i.e. that it cannot be solved by means of *Wissenschaftslogik*), it follows that the grounding is empirical. And this means, from a logical-empiricist vantage point, that the sciences become part of the solution to the “basis problem” in equal measure. If human agents are observers only insofar as (and as much as) any other measuring instrument can be, a theory of measuring instruments is the solution to the basis problem we are seeking. We come to see that our reading of Carnap on protocol sentences in 1932, as prompted by Feyerabend’s “disinfected empiricism”, reveals a Carnap perfectly willing to support a bootstrapping explanation of the empirical basis of the sciences by recourse to the methods and results of empirical science.

## References

Awodey, S., and A. W. Carus. 2006. “Carnap’s Dream: Gödel, Wittgenstein, and Logical, Syntax.” *Synthese* 159 (1): 23–45. doi:10.1007/s11229-006-9066-4.

Butts, Robert E. 1966. “Feyerabend and the Pragmatic Theory of Observation.” *Philoso-*

---

happen to correspond to individual selfs. “Methodological solipsism” simply recognizes this contingent fact. (Counterfactually, if humans possessed a hive mind like the fictional alien race Borg, processing the input collectively, “methodological solipsism” would have to be predicated of the whole population.)

<sup>46</sup>See Uebel (2007, 287–92) for a discussion of Carnap and Popper’s anti-psychologism.

*phy of Science* 33 (4): 383–94. doi:10.2307/186641.

Carnap, Rudolf. 1928. *Der logische Aufbau der Welt*. Berlin-Schlachtensee: Weltkreisverlag.

———. 1932a. “Die physikalische Sprache als Universalsprache der Wissenschaft.” *Erkenntnis* 2: 432–65.

———. 1932b. “Erwiderung auf die vorstehenden Aufsätze von E. Zilsel und K. Duncker.” *Erkenntnis* 3: 177–88.

———. 1932c. “Psychologie in physikalischer Sprache.” *Erkenntnis* 3: 107–42.

———. 1932d. “Über Protokollsätze.” *Erkenntnis* 3 (2-3): 215–28.

———. 1934a. *Logische Syntax der Sprache*. Schriften zur wissenschaftlichen Weltauffassung 8. Wien: Julius Springer.

———. 1934b. *The Unity of Science*. London: Kegan Paul, Trench, Trubner & Co., Ltd.

———. 1936. “Testability and Meaning - Part 1.” *Philosophy of Science* 3 (4): 419–71.

———. 1937. “Testability and Meaning - Part 2.” *Philosophy of Science* 4 (1): 1–40.

———. (1939) 1957. *Foundations of Logic and Mathematics*. Foundations of the Unity of Science 1.3. Chicago, IL: University of Chicago Press.

———. 1959. “Psychology in Physical Language.” In *Logical Positivism*, edited by A. J. Ayer. New York: Free Press.

———. 1987. “On Protocol Sentences.” Translated by Richard Creath and Richard Nollan. *Noûs* 21 (4): 457–70.

———. (1937) 2014. *Logical Syntax of Language*. Routledge.

Carus, A. W. 2007. *Carnap and Twentieth-Century Thought: Explication as Enlightenment*. Cambridge, UK; New York, NY: Cambridge University Press.

Crawford, Sean. 2014. “On the Logical Positivists’ Philosophy of Psychology: Laying a Legend to Rest.” In *New Directions in Philosophy of Science. The Philosophy of*

*Science in a European Perspective Vol. 5*, edited by Maria Carla Galavotti, Dennis Dieks, Wenceslao J. Gonzalez, Stephan Hartmann, Thomas Uebel, and Marcel Weber, 711–26. Springer.

Duhem, Pierre. 1906. *La Théorie Physique: Son Objet, et sa Structure*. Paris: Chevalier & Rivière.

Feest, Uljana. 2017. “Physicalism, Introspection, and Psychophysics: The Carnap/Duncker Exchange.” In *Eppur Si Muove: Doing History and Philosophy of Science with Peter Machamer - A Collection of Essays in Honor of Peter Machamer*, edited by Marcus P. Adams, Zvi Biener, Uljana Feest, and Jacqueline A. Sullivan. The Western Ontario Series in Philosophy of Science. Springer.

Feigl, Herbert. 1958. “The ‘Mental’ and the ‘Physical’.” *Minnesota Studies in the Philosophy of Science* 2: 370–497.

———. 1967. *The Mental and the Physical: The Essay and a Postscript*. University of Minnesota Press.

Feyerabend, Paul K. 1951. “Zur Theorie der Basissätze.” Dissertation zur Erlangung der Doktorgrades an der philosophischen Fakultät der Universität Wien, Wien: Universität Wien.

———. 1954. “Physik und Ontologie.” *Wissenschaft und Weltbild: Monatsschrift für alle Gebiete der Forschung* 7 (11/12): 464–80.

———. 1958. “An Attempt at a Realistic Interpretation of Experience.” *Proceedings of the Aristotelian Society* 58: 143–70. doi:10.2307/4544593.

———. 1962. “Explanation, Reduction, and Empiricism.” In *Scientific Explanation, Space, and Time*, edited by Herbert Feigl and Grover Maxwell, 3:28–97. Minnesota Studies in the Philosophy of Science. Minneapolis: University of Minnesota Press.

———. 1965a. “Problems of Empiricism.” In *Beyond the Edge of Certainty: Essays in Contemporary Science and Philosophy*, edited by Robert G. Colodny, 2:145–260. Uni-

versity of Pittsburgh Series in the Philosophy of Science. Englewood Cliffs (NJ): Prentice-Hall.

———. 1965b. “Reply to Criticism: Comments on Smart, Sellars and Putnam.” In *Proceedings of the Boston Colloquium for the Philosophy of Science 1962-64: In Honor of Philipp Frank*, edited by R. S. Cohen and M. W. Wartofsky, 223–61. Boston Studies in the Philosophy of Science 2. New York: Humanities Press.

———. (1951) 2010. “Die Dogmen des Logischen Empirismus.” In *Vertreibung und Rückkehr der Wissenschaftstheorie: Am Beispiel von Rudolf Carnap und Wolfgang Stegmüller*, edited by Friedrich Stadler, 410–36. Vienna-Berlin-Münster: LIT Verlag.

———. (1954) 2015. “Physics and Ontology (1954).” In *Physics and Philosophy*, edited by Stefano Gattei and Joseph Agassi, 9–24. Cambridge, UK; New York, NY: Cambridge University Press.

———. (1948) 2016. “The Concept of Intelligibility in Modern Physics (1948).” Translated by Daniel Kuby and Eric Oberheim. *Studies in History and Philosophy of Science Part A* 57: 64–66. doi:10.1016/j.shpsa.2015.11.016.

Fraassen, Bas C. Van. 1992. “From Vicious Circle to Infinite Regress, and Back Again.” *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1992 (January): 6–29. doi:10.2307/192822.

Friedl, Johannes. 2013. *Konsequenter Empirismus. Die Entwicklung von Moritz Schlicks Erkenntnistheorie im Wiener Kreis*. Vol. 3. Moritz Schlick Studien. Wien-New York: Springer.

Hanson, Norwood Russell. 1958. *Patterns of Discovery. An Inquiry into the Foundations of Science*. Cambridge: Cambridge University Press.

Hoyningen-Huene, Paul. 1995. “Two Letters of Paul Feyerabend to Thomas S. Kuhn on a Draft of the Structure of Scientific Revolutions.” *Studies in History and Philosophy of Science Part A* 26 (3): 353–87. doi:10.1016/0039-3681(95)00005-8.



Hull, Richard T. 1972. "Feyerabend's Attack on Observation Sentences." *Synthese* 23 (4): 374–99. doi:10.1007/BF00636292.

Kuby, Daniel. 2010. "Paul Feyerabend in Wien 1946–1955: Das Österreichische College und der Kraft-Kreis." In *Auf der Suche nach authentischem Philosophieren. Philosophie in Österreich 1951–2000. Verdrängter Humanismus-verzögerte Aufklärung*, edited by M. Benedikt, R. Knoll, F. Schwediauer, and C. Zehetner. Verdrängter Humanismus - verzögerte Aufklärung: Philosophie in Österreich von 1400 bis heute 6. Wien: WUV.

———. 2016. "Feyerabend's 'the Concept of Intelligibility in Modern Physics' (1948)." *Studies in History and Philosophy of Science Part A* 57 (June): 57–63. doi:10.1016/j.shpsa.2015.11.004.

———. forthcoming. "Feyerabend's Pragmatic Theory of Observation." (In preparation).

Kuby, Daniel, Christoph Limbeck-Lilienau, and Michael Schorner. 2010. "Die Dogmen des Logischen Empirismus von Paul Feyerabend: Editorische Notiz." In *Vertreibung und Rückkehr der Wissenschaftstheorie: Am Beispiel von Rudolf Carnap und Wolfgang Stegmüller*, edited by Friedrich Stadler, 403–9. Vienna-Berlin-Münster: LIT Verlag.

Kuhn, Thomas S. 1962. *The Structure of Scientific Revolutions*. First Edition. Chicago, IL: University of Chicago Press.

Maxwell, Grover. 1962. "The Ontological Status of Theoretical Entities." In *Scientific Explanation, Space, and Time: Minnesota Studies in the Philosophy of Science*, edited by Herbert Feigl and Grover Maxwell, 3:181–92. University of Minnesota Press.

Misak, C. J. 1995. *Verificationism: Its History and Prospects*. London and New York: Routledge.

Morris, Charles William. 1938. *Foundations of the Theory of Signs*. Edited by Otto Neurath. International Encyclopedia of Unified Science 1.2. Chicago: University of Chicago Press.

Oberdan, Thomas. 1990. "Positivism and the Pragmatic Theory of Observation." *PSA*:

*Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1990 (January): 25–37. doi:10.2307/192690.

———. 1993. *Protocols, Truth and Convention*. Amsterdam: Rodopi.

Oberheim, Eric. 2006. *Feyerabend's Philosophy*. Berlin: Walter de Gruyter.

Olegario da Silva, Gilson. 2013. "Carnap and Kuhn on Linguistic Frameworks and Scientific Revolutions." *Manuscrito* 36 (1): 139–90.

Popper, Karl R. 1959. *The Logic of Scientific Discovery*. London: Hutchinson.

Preston, John. 1997. *Feyerabend: Philosophy, Science, and Society*. Cambridge: Polity Press.

Quine, W. V., and Rudolf Carnap. 1991. *Dear Carnap, Dear van: The Quine-Carnap Correspondence and Related Work: Edited and with an Introduction by Richard Creath*. Edited by Richard Creath. Berkeley, Ca.: University of California Press.

Richardson, A. 1997. "Toward a History of Scientific Philosophy." *Perspectives on Science* 5 (3): 418–51.

Schilpp, Paul Arthur. 1963. *The Philosophy of Rudolf Carnap*. La Salle, Ill.: Open Court.

Stadler, Friedrich. 2010. "Paul Feyerabend and the Forgotten 'Third Vienna Circle'." In *Vertreibung, Transformation Und Rückkehr Der Wissenschaftstheorie: Am Beispiel von Rudolf Carnap Und Wolfgang Stegmüller*, edited by Friedrich Stadler, 169–87. Vienna-Berlin-Münster: LIT Verlag.

Townsend, Burke. 1970. "Feyerabend's Pragmatic Theory of Observation and the Comparability of Alternative Theories." *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1970 (January): 202–11. doi:10.2307/495764.

Uebel, Thomas E. 2007. *Empiricism at the Crossroads: The Vienna Circle's Protocol-Sentence Debate*. Chicago: Open Court.

———. 2013. "Pragmatics in Carnap and Morris and the Bipartite Metatheory Conception." *Erkenntnis* 78 (3): 523–46. doi:10.1007/s10670-011-9352-5.

Yuann, J. J. 2007. "A Naturalistic Approach to Scientific Methodology: A Comparative Study of O. Neurath and P. Feyerabend." In *Naturalized Epistemology and Philosophy of Science*, edited by Michael Mi Chienkuo, 171–96. New York: Rodopi NY.