Beyond Representationalism: Reframing Translation Processes through the Enactivist ABC Framework

[*1]Michael Carl, [2]Xinyue Ren

ORCID: 0000-0002-2815-0292, 0000-0001-6239-7195

[1]Kent State University, [2]City University of Hong Kong

Abstract

This paper addresses the persistent representational bias in Cognitive Translation Studies which typically models translation as higher-order inferential communication grounded in metarepresentational mindreading. Drawing on Relevance Realization (RR, Vervaeke et al. 2012) and the ABC framework of enacted cognition (Carl 2025b), we argue that translation can be adequately understood in non-representational terms, as an embodied, affectively modulated, and culturally scaffolded process of sense-making. Re-examining Gutt's (2005) distinction between S-mode and I-mode translation through an enactivist lens, we show that these modes need not presuppose contentful mental representations but can be interpreted as distinct regimes of skilled engagement and reflective recalibration within a dynamically structured activity space.

We propose that relevance is not assessed over mental representations but dynamically enacted through the selective organization of translational affordances across time. We conceptualize this view as an extension to Relevance Theory (RT, Sperber & Wilson 1986; Gutt 1991/2000) by grounding relevance in resemblance in organism-environment coupling rather than internal symbol manipulation. The account preserves RT's normative orientation while offering a unified explanation of translation cognition that naturally accommodates priming, affect, intuitive judgment, and cultural attunement.

Keywords: Relevance Theory; Relevance Realization; S-mode and I-mode Translation; ABC Framework of Enacted Cognition

1. **Introduction**

Relevance Theory (RT, Sperber & Wilson 1986; Gutt 1991/2000) enjoys increasing popularity in Translation Studies. RT makes a distinction between descriptive and interpretive language use. Descriptive language use aims to linguistically represent an assumed state of affairs in the external world, while interpretive language use aims at simulating the communicative function or an assumed mental state behind an utterance. Gutt (1991/2000) draws on RT to define translation as an instance of *interpretive language use*, whereby the translator seeks to establish interpretive resemblance between a source text and the target text, understood not as propositional equivalence but as similarity in contextual effect.

In both modes of language use (descriptive and interpretive), RT endorses a modern, inferential representationalism, compatible with a modular cognitive architecture, in which the mind is assumed to reason by transforming representational states according to inferential principles. Even though RT rejects the classical Fodorian-style code-based models of meaning transmission, in which communication involves encoding mental content into language, transmitting it, and decoding it on the listener's end, RT is strongly anchored in mental representationalism.

Thus, interpretive resemblance is assumed to be achieved through mind-reading: inferring the speaker's beliefs and intentions, and by reconstructing the original communicative goal. RT, thereby, presupposes a Higher-Order Theory of Mind (HOToM, Sperber 1997; Origgi & Sperber 2000), which posits that understanding others requires forming metarepresentations of their mental states (Wilson 2000). HOToM, as well as RT, assume that mental states become conscious through higher-order thoughts about lower-level states; cognition is defined by the manipulation and nesting of contentful mental representations (Berger & Brown 2022).

Our reading of Gutt's (1991/2000, 2004, 2005) conception of translation as interpretive language use hints at an extension of modern representationalism. Rather than prioritizing the preservation of truth-conditional content of nested propositions, Gutt emphasizes the normative and functional dimension of translation. Translation establishes, according to him, a relation of resemblance between "bodies of thought," (Gutt 2004) i.e., culturally and pragmatically situated intentions, rather than propositional truths. In a further development of this idea, Gutt (2005) distinguishes between two translation modes, a stimulus mode (S-mode) and an interpretive

mode (I-mode). While S-mode translation is grounded in the transfer of linguistic/form-based signals (i.e., communicative clues), I-mode translation focuses on interpretive resemblance, which can involve norms, emotions, and cultural embeddings—none of which are, in our view, exhaustively representational contentful.

The present paper develops an enactivist reading of Gutt's (2005) S-mode and I-mode translation. It alludes to radical enactivism (Hutto & Myin 2012, 2017; Kirchhoff & Kiverstein 2019), rejecting that cognition requires contentful internal representations at all. In the enactivist view, translation can be seen as emerging from embodied, affective, and context-sensitive interactions with the environment. Mental states, including those about oneself and others, need not be constructed via propositional inference. Rather than representing, they serve to enact embodied engagement, emotional attunement, and pragmatic responsiveness. Higher-order states are then scaffolded by cultural practices, narrative engagement, and emotional resonance rather than by abstract metarepresentational processing.

2. **Representationalism**

The classical representational view of the mind argues that mental processes involve internal representations, structured, contentful symbols, or analogues that stand for things in the world. On this view, cognition operates by manipulating these representations according to formal rules, much like a computer processes symbolic information. This framework has been central to cognitive science since the mid-20th century.

In modern representationalism, the existence and nature of the mind must be inferred (Berger & Brown 2022). Theory of Mind (ToM, Premack & Woodruff 1978) refers to our ability to attribute mental states, such as beliefs, desires, intentions, and emotions, to others and to ourselves. It is usually assumed that others have minds like our own, and thus ToM is the capacity to understand that other people have their own beliefs, desires, and intentions, which can differ from ours. We use this understanding to explain and predict their behavior. However, ToM does not give us direct access to others' mental states, but it helps us model them through inference or simulation. This ability, it is claimed, allows us to reason about what others think,

feel, or know, and it enables us to reason about minds, which is believed to be essential for translation.

2.1. ***Higher-Order ToM (HOToM)***

Higher-Order ToM (HOToM) adds a degree and complexity to ToM (Rosenthal 2005). While ToM is the foundational ability to attribute first-order beliefs to someone else (e.g., "He believes that X"), HOToM requires representing second (or higher) order beliefs, beliefs about beliefs, such as "Anne thinks that Sally thinks the marble is in the basket." Whereas ToM assigns a speaker the ability to attribute mental states (beliefs, desires, intentions, emotions) to others, and to understand that these states may be different from their own, HOToM attributes mental states about other mental states, i.e., meta-mental state attribution.

(HO)ToM can be conscious or unconscious. Conscious (HO)ToM is the deliberate, reflective awareness of another's mental states, or an awareness of what someone thinks about what someone else thinks. It explains counterfactual reasoning, narrative construction, perspective-shifting or complex judgments like deception, sarcasm, or irony. It supports metapragmatic acts like clarification, justification, or apology. Accordingly in translation, conscious (HO)ToM supports grasping nuances like implied meanings, tone, or indirect speech acts.

2.2. ***Hot and Cold ToM***

According to some researchers (e.g., Prinz 2004; Sturm 2020), there are two kinds of representations: cognitive (concerning mental states, beliefs, thoughts, and intentions) and affective (concerning arousal, feelings and emotions). A distinction between hot (affective) and cold (cognitive) ToM was introduced in the developmental and social cognitive neuroscience literature (e.g., Draperi et al. 2022), particularly by researchers interested in the emotional versus cognitive components of understanding others' minds. However, we will use here the terms hot/cold ToM, rather than equivalent affective/cognitive.

Cold ToM refers to the capacity to attribute mental states to others (e.g., beliefs, intentions, knowledge) in a detached, inferential way, often relying on reasoning and perspective-taking rather than emotional engagement. For example, the sentence "She believes it's going to rain, so she's taking an umbrella" involves reasoning about the mental states of others.

Cold ToM processes can be automatic, fast, and nonconscious (Butterfill & Apperly 2013). Cold ToM has been observed in infants as young as 7–9 months who show sensitivity to others' beliefs in nonverbal paradigms — well before they possess full-blown metacognitive or linguistic capacities, suggesting nonconscious processing (Hirshkowitz & Rutherford 2021). But also adults make rapid social inferences (e.g., recognizing someone is unaware of something) without explicitly reasoning — supporting the idea that cold ToM can operate below the threshold of awareness (McCrea 2010).

Hot ToM refers to a facet of ToM that involves understanding and attributing emotions and desires to others, particularly in situations involving affect and motivation. It contrasts with cold ToM, which focuses on understanding others' beliefs and thoughts in a more detached, cognitive manner. An example of hot ToM would be the capacity to feel sadness when reading a tragic character's monologue — not just knowing they are sad but experiencing (simulating) that sadness. Draperi et al. (2022) observe that "hot ToM abilities [are] developed more rapidly than cool ToM" in children aged between 3.5 to 6.5 years. Hot ToM has a direct relation with qualia[1] as it involves emotionally simulating or resonating with others' feelings, reproducing or mirroring another's qualia within your own affective system, via simulation or emotional contagion.

2.3. ***Relevance Theory (RT)***

RT (Sperber & Wilson 1986) is a theory of pragmatics and communication that aims at explaining how hearers infer a speaker's meaning, based on the notion of *cognitive relevance*, i.e., maximizing informational gain while minimizing effort. As mentioned above, RT rejects the

[1] The subjective, qualitative, or introspectively accessible properties of mental states, particularly the "what it is like" aspect of conscious experience.

conduit model of communication and refines the notion of representation as the basis of an interpretive process, in which relevance operates over representations. This makes RT a form of modern representationalism (Berger & Brown 2022). Representations are understood through their functional roles rather than as static mental images. As for (HO)ToM, also according to RT, a hearer constructs a mental representation of the speaker's intended thought. Thoughts are here propositions, representations with conceptual structure that are amenable to truth-evaluation in the case of explicatures, or attitude attribution in the case of implicatures (Sturm 2020).

Meanings, in this view, are content-bearing mental structures, and metarepresentations are the core commitment by which minds store and operate on representations of other minds. According to Sturm (2020, 46), "The implicit notion of metarepresentation refers to the text as the representation of reality."

RT draws on the idea that communication is ostensive-inferential, that is, communication depends on mutual recognition of communicative intentions. RT assumes thereby that for something to be communicated successfully, it must be mentally represented, accessible, and relevant to the audience. Thus, communicators engage in ostensive communication as they intentionally and consciously make their mental states manifest to others.

According to RT, cognition and communication are grounded in the capacity of representing and metarepresenting. Cognitive systems, as Mercier & Sperber (2017, 3) put it, "are characterized by their ability to construct and process mental representations." Representations in this context refer to internal models that stand in for states of affairs in the world. Metarepresentations, or second-order representations, go a step further: they are mental representations of mental representations — for example, my belief that "John believes that it will rain" reflects my cognitive capacity to represent another's belief.

2.3.1. *HOToM and RT of Translation*. Drawing on Gutt (2004, 2005), Sturm (2017) posits that translation is not merely a linguistic exercise, but that translation engages HOToMs: Translators must attribute and infer mental states (beliefs, intentions, emotions) to both the original author and the target audience. Sturm (among others, e.g. Gallai 2022) highlights that translators need to model and represent the minds of others during the translation process. This involves anticipating how the target audience will interpret the translated text, which opens the possibility for the translator to adjust the translation and to maintain the intended meaning and effect.

RT conceptualizes "translation [as] a transformation of metarepresentations" (Sturm 2017), where metarepresentations are second (or higher) order representations, representations of representations of others, i.e., HOToMs. Accordingly, translation becomes "a simple chain of text production and re-production between author, translator and audience [where] Metacognition [and metarepresentation] is a central feature of this process." (Sturm 2017, 421) "Metarepresentation provides a simple, yet powerful model of translation as a special form of inferential communication." (ibid.)

2.3.2. *Hot/Cold ToM in RT.* Communication in the RT framework is considered a kind of 'mind-reading' or metarepresentation: I understand what someone means only if I can depict what they believe I know, and why they are saying what they say in the given context. When a friend tells me: "It's cold in here," I understand that the friend wants me to realize that s/he wants me to do something, (e.g., close the window). This second-order ToM ("S/he believes I can infer what s/he wants based on what s/he said") is an example of (cold) HOToM.

The interaction of hot and cold ToM with HOToM becomes important in more affectively charged situations, such as, for instance, when a colleague tells me, with a smile but a flat voice:

"Oh great, another meeting ..."

Hot and cold ToM interact here with HOToM in various ways:

1. Cold ToM: I infer: "S/he believes meetings are bad."

2. Hot ToM: I feel: "S/he is annoyed or frustrated."

3. HOToM: I realize: "S/he wants me to recognize that s/he's being sarcastic."

4. Hot ToM + HOToM: I realize: “S/he wants me to recognize that s/he’s being sarcastic and join their emotional stance.”

which may then lead to a respond with empathy, humor, or irony, as, for instance: “Yeah, just what we needed, right?” Here, cold ToM allows me to infer their beliefs, intentions, or knowledge, in this case to infer the intentional structure behind the utterance (“S/he believes the meeting is bad.”). Cold HOToM is necessary to attribute nested mental states, e.g., “S/he believes I know s/he believes this.”

Hot ToM addresses the emotion behind the sarcasm which enables an emotionally attuned inference about feelings and affective states. Specifically, it adds emotional color to HOToM: it picks up on tone, intonation, facial expression, or body language to grasp the feeling behind the sarcasm, possibly frustration or resignation, etc. Thus, hot ToM + HOToM helps us to understand “S/he wants me to know that s/he is annoyed and wants me to feel that with them.” Empathy is here explained through hot ToM, a cognitive mechanism where emotional understanding is treated as an interpretation of the communicator’s intended meaning. It is considered an inferential process by which listeners infer affective states according to the principle of relevance. This mechanism is, according to RT, not only crucial for understanding non-literal speech acts like sarcasm, but also for translating (Gallai 2022). Inferential HoT mechanisms allow translators to grasp what an author thinks or feels, as well as what the author expects its audience to understand about what they think or feel. Thus, RT aligns with ToM, where understanding is conceptualized as the evaluation of nested mental states:

- cold ToM enables a translator to:
    - Recognize intentions, beliefs, and speech acts (e.g., a promise, a warning).
    - Maintain logical coherence and referential clarity.
    - Reconstruct implicatures and presuppositions (Gricean inferences).
    - Engage in monitoring and revision phases
      (e.g., the translator consciously evaluates coherence and relevance).
- hot ToM enables a translator to:
    - Understand the emotional perspective of the speaker/author.
    - Emotionally attune to tone, mood, irony, satire, or pathos.
    - Anticipate the emotional effect on the target audience and modulate tone accordingly.

- Immersive, in flow-like drafting and embodied empathy
  (e.g., entering the emotional stance of a fictional narrator).

Following Sperber and Wilson (1986), Gutt (2004, 13) mentions that "metarepresentational capacities must be the core component of translation competence," where metarepresentation is "nothing more than what the translator believes regarding the relationship between the author and his/her audience" (Sturm 2020, 40).

However, in what follows, we propose an enactive interpretation in which representational assumptions are not mandatory for RT (Vervaeke et al. 2012). Post-representational (or non-representational) positions have been suggested that align with the core tenets of RT, in which representations are not basic in cognition but may emerge through processes of stabilization. Cognition, correspondingly, is not primarily about constructing accurate internal representations of the world, but about dynamically tuning the agent–world coupling so that certain affordances become salient.

3. **Relevance Realization (RR)**

From a situated, action-oriented point of view, cognition first of all supports the dynamic coupling between organism and environment. Meaning and understanding do not primarily arise from the manipulation of internal representations, but they emerge through interaction, affordances, and ongoing sensorimotor engagement.

On this view, metacognition can be reconceptualized as the capacity to participate in, and skillfully respond to, the intentions and actions of others within embodied and socially structured contexts, rather than as the manipulation of internal symbolic structures. Enactivist theories of mind (e.g., Varela et al. 1991; Gallagher 2001, 2017) accordingly suggest that understanding others' mental states does not require forming abstract beliefs about their beliefs. Instead, such understanding is achieved through coordinated social practices, affective attunement, and pragmatic responsiveness within shared situations.

This perspective is not incompatible with the core commitments of RT, and in particular the principle of relevance, according to which cognition is geared toward the optimization of relevance. In communicative contexts, optimal relevance is defined as achieving an appropriate balance between maximizing cognitive effects and minimizing processing effort. Thus, Gutt (1991/2000) claims that "the relevance framework provides all the theoretical concepts necessary for translation". Relevance thus functions as a guiding constraint that steers cognition toward outcomes yielding significant cognitive benefit, such as novel or actionable information, at minimal cost.

In its standard formulation, RT assumes that relevance is assessed over a field of available mental representations. Cognitive effects are taken to arise through the inferential processing of these representations, and relevance is defined relative to the effort required to derive such effects. While the principle of relevance therefore presupposes representational content, RT does not provide a general account of how the underlying representational structures themselves emerge (cf. Vervaeke et al. 2012). From a broader cognitive-theoretical perspective, this leaves open the question of how relevance can be operative prior to, or independent of, fully articulated representations.

We propose a non-standard reading of RT that suggests a reversal of this explanatory order, where representations are the possible *outcomes* of relevance-sensitive processes rather than their starting point. In this vein, Vervaeke et al. (2012) introduces the notion of *relevance realization* (RR), drawing on a dynamical systems perspective to explain how cognitive systems differentiate signal from noise and thereby bring about stable patterns of significance. RR is a pre-representational process of self-organization through which certain aspects of the agent–environment interaction become salient and actionable, prior to contentful representation (Jaeger et al. 2024). Representations are thus better understood as relatively stabilized frames or constraints that emerge from a dynamic, non-representational "grip" on a landscape of affordances (Rietveld & Kiverstein 2014).

Vervaeke's notion of RR[2] explicitly undercuts classical (and modern) representationalism, even though it does not entirely dispense with representational vocabulary. Representations,

---

[2] See https://www.meaningcrisis.co/all-transcripts/ for a highly interesting lecture series on Vervaeke's notion of RR.

where they arise, are secondary, emergent, and pragmatically constrained rather than foundational explanatory posits. In this view, representational structures arise only through the stabilization of relevance relations in social and normative practices (e.g., De Jaegher & Di Paolo 2007; Hutto & Myin 2012).

RR thus provides a non-representational (or post-representational) grounding for RT that is fully compatible with enactivist and embodied approaches. It suggests that the principle of relevance (trade-off between processing effort and cognitive effects) can be reinterpreted as norm-sensitive constraints that govern how emergent structures of significance stabilize and guide action. RT, in this non-standard view, becomes less a theory of representational content and more a theory of the functional regularities that regulate how meaning, salience, and understanding are enacted in situated cognitive activity.

3.1. ***RR and Higher-order Awareness***

From an RR point of view, higher-order awareness of our (or another's) mental state is not first of all a representation of a representation, but a form of embodied access and skillful interaction with one's own (or another's) cognitive processes. Higher-order awareness is the capacity to dynamically bring certain cognitive processes into salience because they matter for current and future action.

For instance, consider a situation where a visitor in my house says, "It's cold in here," and I respond by closing the window. Within standard RT, this process would be framed as an instance of inferential communication which presumes to meta-represent others' (the guest's) mental states and to reason about them. This seemingly simple exchange involves several steps:

1. First, I understand the *literal meaning* of the utterance and their belief that the room is cold.
2. Then, I understand the *speaker's intention*: their desire to change the temperature.
3. Finally, I infer the communicative intent: The speaker wants me to realize that s/he wants me to do something (close the window).

Crucially, this requires me to infer what they believe I will understand and how they expect me to act. The inference third step is a case of Higher-Order (i.e., second-order) HOToM, in which I attribute beliefs about beliefs and intentions.

From the RR perspective, this interaction can be described in terms of the dynamic modulation of salience within an unfolding agent–environment system. The utterance "It's cold in here" functions as a perturbation that reorganizes the relevance landscape of the situation. Rather than triggering a sequence of meta-representational inferences about the speaker's beliefs and beliefs-about-beliefs, the utterance selectively enhances the salience of certain affordances, most notably the open window as a site of possible action.

Higher-order awareness, on this account, consists in the listener's capacity to become sensitive to how this perturbation matters for ongoing and future interaction. The listener does not need to represent the speaker's mental state as an object of thought; instead, the listener is dynamically attuned to the practical implications of the situation, as shaped by shared norms of hospitality, bodily comfort, and cooperative interaction.

In this way, RR explains how agents are able to settle on a determinate interpretation and course of action without evaluating an infinite space of possibilities, overcoming the 'radical underdeterminacy' hypothesis (Gutt 2005). Goals and norms function as constraints that shape the dynamics of relevance, allowing certain patterns to stabilize while others fade into irrelevance. Representations, where they arise, are therefore best understood as compressed, reusable summaries of past relevance realizations — tools for coordination that depend on, but do not replace, the underlying non-representational dynamics. In this sense, RR provides a non-representational account of how communicative understanding can be both flexible and norm-governed, while remaining grounded in embodied interaction rather than meta-representational inference.

3.2. ***Translation and RR***

As mentioned above, RT distinguishes between descriptive language use and interpretive language use. The interpretive language use stipulates that an utterance is evaluated by the degree to which it bears resemblance to another utterance or thought, rather than with respect to its truth-conditional correspondence to a state of affairs. Gutt (1991/2000) argues that translators establish interlingual interpretive resemblance by recreating its relevance for a target audience situated within a different cultural and/or normative environment. What is preserved is not

propositional equivalence but a structurally similar pattern of pragmatic effects. Translation, in this sense, is not straightforwardly truth-evaluable, but a norm-sensitive, contextually embedded practice.

From the RR perspective, interpretive resemblance can be understood as the dynamic alignment of structural relations rather than as a comparison between pre-existing representations. The translator's task is not to match contents, but to become selectively attuned to those features of the source text that matter for action, understanding, and coordination in the target context. Precision weighting plays a crucial role in this process: linguistic, cultural, and situational cues are differentially weighted so that they allow for sustaining communicative viability. Through this ongoing modulation of precision, certain structural similarities between source and target utterances are amplified, while others are down-weighted or ignored.

Initially, these similarities are enacted rather than represented. They take the form of biased responsiveness to affordances, i.e., patterns of salience that guide attention, affect, and action without yet constituting determinate content. However, as relevance-guided engagement is repeated and socially ratified, these patterns of structural similarity begin to stabilize. Over time, such stabilization allows resemblance relations to crystallize into representationally tractable structures: reusable, nameable, and normatively constrained ways of understanding what a translation "means" or "does." On this view, representational content is not the precondition of interpretive resemblance, but its sedimented outcome.

This account avoids the representationalist claim that interpretive resemblance must be grounded in shared mental content or explicit meta-representational inference. While standard RT locates resemblance within ostensive–inferential communication governed by mutual cognitive environments, an RR-based interpretation locates it within norm-regulated dynamics of relevance that operate prior to, and beneath, explicit representation. Norms and goals do not function primarily as represented rules, but as constraints that shape the precision landscape of the task, allowing agents to settle on viable interpretations without canvassing an infinite space of possibilities.

From an enactive perspective, interpretive resemblance is thus enacted through affectively modulated, sensorimotorly organized, and socially scaffolded engagement with linguistic and cultural affordances. As Robinson (2023) notes, norms exert much of their force

through intuition, habit, and affective orientation, shaping how uncertainty is tolerated and how future action is anticipated. Translation, in this view, is best understood as a form of skilled coordination and embodied sense-making (Rietveld & Kiverstein 2014, Kiverstein & Rietveld 2021), in which relevance realization guides the emergence, and eventual stabilization, of representational structures, rather than presupposing them.

4. **Emerging Representations**

According to Gutt (2004), translation is an act of communication that is concerned with another communication. He explains that "any act of communication concerned with another act of communication…can aim at providing information about either of its two key elements: the stimulus used, or the interpretation intended in the original act" (Gutt 2005, 32).

Translation can thus focus on either the stimulus ("what is said") or the interpretation ("what is meant"). That is, a translator can operate in the S-mode (stimulus-mode), if the translation reproduces aspects of the original expression, its "communicative clues", i.e., form, structure, or surface features. Otherwise, if the translator focuses on the intended meaning or interpretation, s/he resorts to the I-mode (interpretive mode). Gutt (2005, 35) clarifies that "the stimulus is the perceptible evidence", while the "the intended meaning is the thoughts of the communicator."

In the S-mode, a translator re-uses the original utterance, assuming that duplicating formal characteristics (or a close equivalent) in another language can preserve the intended meaning in which the translator "need not necessarily have understood the intended interpretation of the original at all" (Gutt 2005, 38)[3]. In that case, the translation reproduces the explicatures of linguistic surface features of the source text so that the target audience can reconstruct the intended meaning. However, if representation involves an intentional, cognitive

---

[3] Gutt maintains that interlingual S-mode is, actually, a hybrid of the two basic modes, which he calls *the interpretation-oriented s-mode of higher-order communication across language boundaries*.

act of standing in for an object, concept, or state of affairs, it presupposes a level of understanding which is not granted in S-mode translation.

This suggests that S-mode translation operates on pre-representational interlingual structures. While S-mode translation does not automatically and truthfully describe a state of affairs in another language, Szpak et al. (2025, 71) point out that reproducing the linguistic stimulus of an utterance that was originally used descriptively *can* result in a description of the state of affairs, provided the ST communicative clues are replicated in a faithful way and the target audience is able to recover that same descriptive interpretation. In this case, S-mode translation copies the original's descriptive use, and can be described as truth-oriented, not because the translator asserts the truth of an original utterance, but because the translator enables the target audience to recover the descriptive use that made up the original source utterance. However, S-mode translation is different from descriptive language use. In descriptive language use, which, according to Gutt (2005), is not translation proper, the translator would willfully and intentionally describe a state of affairs in another language, even without necessary reference to the source.

4.1. ***Priming and S-mode Translation***

For Gutt (2005, 40), S-mode translation "involves the use of another token [i.e., the TL item] (reproduction, replica) of the original stimulus [a SL prime] … by the fact that [they] share … properties." The mental processes underlying such stimulus replication can be understood in terms of *priming*, that is, a form of implicit memory in which exposure to one stimulus influences subsequent processing in a related domain, based on resemblance.

In translation, priming does not consist in the direct elicitation of a specific target-language item as a response to a source-language word. Rather, source-text input modulates the translator's subsequent processing dynamics, such as attentional allocation, activation patterns, and action readiness, thereby shaping how later translation operations unfold. In interlingual language use, "S-mode stimulus replication is now in terms of shared communicative clues" (Gutt 2005, 43). The effect of priming in translation lies in the facilitation and stabilization of the

processes from which target-language expressions emerge, rather than in the transmission or reproduction of linguistic items themselves.

Similarly, Schaeffer and colleagues (Schaeffer & Carl 2013; Carl 2025a) suggest that translation is driven by priming (perceptual, structural, semantic, affective, and normative priming). These primes can be understood as functioning like priors for successive translational actions (see also Robinson 2022). In an enactivist framework (e.g., Bruineberg 2017; Kirchhoff & Kiverstein 2019; Ramstead et al. 2019; Rietveld & Kiverstein 2014) priors have been conceptualized as skillful dispositions or embodied expectations rather than (classical) representations. On this view, priors are inseparable from sensorimotor dynamics and cannot be individuated independently of action contexts. For instance, Kiverstein and Rietveld (2018) propose an enactivist reading of predictive processing where priors encode tendencies for action that help organisms anticipate environmental interactions but, according to this interpretation, priors do not represent the environment as semantic content. In the predictive processing framework, priors are continuously updated through perception-action loops that minimize prediction error, operationalized via the Kullback-Leibler (KL) divergence. Sensorimotor enactivists argue that the KL divergence measures informational differences between probability distributions, but facts about covariance do not entail facts about representation.

4.2. ***Stabilization of Representational Content***

Numerous scholars have argued that similarity is neither a sufficient nor a necessary condition for representation-hood (e.g., Gładziejewski 2016; Lee & Calder 2022; Carl 2025b). Resemblance alone cannot ground representational content, since similarity is inherently unconstrained and context-dependent; as Lee and Calder (2022) put it, “structural similarity is insufficient for content.” If X were to represent Y solely in virtue of resembling Y, one would need to specify which aspects of resemblance are relevant and why. Yet this requirement generates a vicious cycle: resemblance can ground representation only if what is relevant is already fixed, but fixing relevance itself presupposes representational content. Consequently, resemblance ends up presupposing the very representational relation it was meant to explain. This leads Vervaeke et al. (2012, 96) to conclude that “the whole account is circular in nature.”

However, relevance is not only central in communication, but in cognition in general: Jaeger et al. (2024, 2) make it clear that "The ability to solve the problem of relevance is a necessary condition and the defining criterion for making sense of a large world." As a way out, Vervaeke et al. (2012, 87) suggest a "theory that proposes self-organizing mechanisms for relevance realization [which] will be able to dissolve the threat of vacuous or cyclic explanations in cognitive science" and argue that "relevance realization [RR] is going to be relative to the interests and goals of an organism" (ibid., 88). In other words, RR breaks the vicious loop by positing a bootstrapping process in which relevance assessment is a *pre-representational* activity grounded in non-representational structural relations. Jaeger et al. (2024, 15) further amounts RR to the "collective co-constitution of the intrinsic goals, repertoires of action, and affordance landscapes of an organism environment system" and a "collapse and reconstruction of all three sets — affordances, goals, and actions — ultimately committing the agent to a particular pursuit" (ibid., 14).

On this view, structural similarities initially function not as representational vehicles but as constraints on successful coordination and regulation. Through cycles of consolidation and stabilization, certain patterns of interaction become entrenched as normatively salient, eventually enabling representational claims to be made *on top of* these stabilized practices rather than serving as their foundation. Representation thus emerges as a socially and normatively articulated achievement, scaffolded by but not reducible to prior non-representational structural coupling.

A translator might then state that a target text X represents a translation of a source text Y, thereby recruiting similarity considerations of various kinds. But this is a post-hoc assessment regarding the translation product; representational status is conventionally established through *claim-making practices* (Hutto & Myin 2012), in which an agent intentionally asserts that X represents Y. However, this claim-making ability already evolves during the translation production.

4.3. *Affective Awareness and I-Mode Translation*

S-mode translation can thus be characterized as non-representational: the translator does not construct higher-order mental representations of the speaker's communicative intentions, nor does S-mode require explicit inferential simulation of mental states or a reflective representation of interlingual resemblance. Translation proceeds through the reproduction of the original stimulus, relying on shared communicative cues rather than on explicit metarepresentational reasoning.

This characterization, however, raises an apparent difficulty. If the translator does not actively consider pragmatic and cultural context, if no higher-order representations are generated, how can they know whether the same interpretation will, or even can, be recovered by the target audience? How can an S-mode translation be faithful if the translator does not verify that the reproduced stimulus actually leads to a resemblance of interpretation in the target context?

One way to resolve this tension is to assume that contextual and affective awareness actively shape what counts as a need for verification in the first place. As Gutt (2005) emphasizes, since the interpretation of communicative clues is context-dependent, "the key factor for discovering the originally intended meaning in the S-mode of higher-order communication is the accessibility of the original context." This feature becomes particularly salient in I-mode translation, where the translator seeks access not only to the speaker's communicative intention but also to the felt appropriateness, tension, or coherence of context-based interpretations.

In this sense, I-mode involves not only reflective and reportable processing, but also a form of affectively informed awareness: translators monitor their understanding through subtle experiential cues such as uncertainty, fluency, resonance, or dissonance, which function as indicators of interpretive adequacy (Robinson 2022). These affective signals are not external to reasoning; rather, they guide attention, trigger reanalysis, and support the justification of translational choices (e.g., "this formulation feels more consistent with the tone and therefore aligns with norm X"). Meaning is thus reconstructed through processes that are simultaneously constrained by cultural conventions, genre expectations, pragmatic implicatures, and norms of coherence, and modulated by affective evaluations that render these constraints experientially accessible.

I-mode translation does not amount to classical representational transfer either. Rather than manipulating internal symbols, in our view, the translator engages here in felt, participatory and context-sensitive sense-making, interacting with the speaker, the anticipated audience, and the communicative situation. The translated utterance resembles the source not by preserving propositional content, but by fulfilling a similar interpretive function within a communicative act. As Gutt (2005) notes, this often entails modifying content to fit the target context.

From this perspective, effective translation typically involves a dynamic interplay between S-mode and I-mode rather than a strict dichotomy. But, as Hoza (2016, 62 ff.) suggests in his discussion of "in-the-zone" experiences, translators often operate in a fundamentally action-oriented, non-representational state of mind.

4.4. ***S-Mode and I-Mode in Metaphor Translation***

The oscillatory shift between S- and I-mode becomes especially visible in the translation of metaphors, not as vehicles of propositional meaning but as embodied enactments of affective and sensorimotor dynamics. Consider the sentence "He stormed into the discussion." From an enactivist perspective, the translator does not reconstruct hidden mental content behind the metaphor. Instead, the expression conveys a pattern of abruptness, tension, and force, which primes the translator to respond with different degrees of embodied attunement. In an initial S-mode engagement, the translator's behavior is modulated primarily by perceptual and semantic priming, guiding immediate sensorimotor coupling with the surface form. The verb *stormed* recruits embodied associations of rapid, forceful movement, and the translator may simply extend this dynamic into the target language, producing a Chinese translation such as

(1) *他 如 风暴 般 闯入 讨论*。

tā rú fēng-bào bān chuǎng-rù tǎo-lùn

3sg like wind-storm like break-enter discuss-debate

[lit. He, like a storm, burst into the discussion.]

Here, translation arises from relatively automatic responsiveness to the linguistic stimulus, rather than from sustained, reflective interpretation. Such responsiveness is not

modeled as the manipulation of representations; it can instead be understood as skilled, habitual sensitivity to the expressive contours of the linguistic elements. At an initial stage of engagement, the translator remains closely coupled to the surface stimulus. The metaphorical use of *stormed* activates embodied associations of forceful movement, leading to a rendering that preserves structural resemblance.

As the translator shifts into I-mode, their engagement becomes shaped by procedural and affective priming. In this mode, the translator coordinates with the interpersonal affordances evoked by the metaphor. The forcefulness of *stormed* is experienced as a socially familiar pattern of intrusion or assertive entry. This may yield translations such as

(2) *他 突然 加入 了 讨论，态度 强硬。*

tā túrán jiā-rù le tǎo-lùn tài-dù qiáng-yìng

3sg suddenly add-enter pfv discuss-debate attitude firm-hard

[lit. He suddenly joined the discussion, with a forceful attitude.]

In (2), the translator's adjustment can be described as responsiveness to the situational pull of the scene. As the translation process unfolds, contextual and interpersonal considerations become increasingly salient. The translator may recalibrate the emerging target text in light of anticipated audience uptake, resulting in a more interpretive formulation.

At deeper levels of I-mode, the translator may resonate affectively with the interactional tension and enact this resonance through an emotionally charged expression like

(3) *他 气势汹汹 地 闯进 了 讨论 。*

tā qìshì-xiōngxiōng de chuǎng-jìn le tǎo-lùn

3sg imposing-fierce adv break-enter pfv discuss-debate

[lit. He burst into the discussion in a fierce, imposing manner.]

The three translations above may emerge from successive revisions of the unfolding engagement with the source text. They are intended here to illustrate how translational sense-

making can reorganize progressively over time as salience relations, affective orientation, and contextual constraints are recalibrated during the translation process.

From an RR perspective, these successive revisions can be understood as iterative reorganizations of the relevance landscape. Each stage involves a collapse and reconstruction of salience relations among affordances, goals, and anticipated effects (Jaeger et al. 2024). They describe dynamic re-weighting of translation strategies within an evolving field of engagement which would otherwise appear to be a deeper inferential reconstruction.

The progression from (1) to (3) thus illustrates how S- and I-modes are not discrete strategies but fluid regimes of engagement within a single unfolding translation episode. The translator's task may therefore be characterized as navigation within a shifting landscape of affordances (Kiverstein & Rietveld 2021; Rietveld & Kiverstein 2014; Rietveld et al. 2018), structured by affective cues, sensorimotor patterns, and culturally sedimented norms. In this sense, interpretive resemblance is enacted through sustained grip on the evolving situation and skillful responsiveness to what Kiverstein and Rietveld (2021) term "enlanguaged affordances," namely action possibilities structured by linguistic and cultural practices. The ongoing re-organization of affective, sensorimotor, and contextual dispositions stabilizes a sense of meaning, gradually resulting in what we conventionally refer to as representations.

5. **The ABC Framework of Enacted Cognition**

The ABC framework (Carl 2025b; Carl et al. 2025) offers an extension of RT. Instead of treating cognition as symbolic processing inside the skull, the ABC framework conceives sense-making as a temporally unfolding, situated process in which brain, body, tools, and cultural norms co-constitute what an agent can perceive, anticipate, and do. Cognition is understood as the dynamically coordinated interplay between three layers of activity: Affective (A), Behavioral–sensorimotor (B), and Cognitive–conceptual (C).

The ABC framework models how relevance-sensitive translation processes unfold across temporally layered dynamics. It rests on three commitments. First, the framework is non-representational at its base level. Translation is understood as affectively and behaviorally

modulated coupling through ongoing engagement with linguistic and cultural affordances. The A- and B-layers do not encode contentful internal representations but express embodied dispositions and precision-weighted tendencies that regulate action. The framework predicts the stabilization of salience structures and the emergence of representations at the C-layer. Representations are not taken as explanatory primitives but as socially and normatively stabilized outcomes of RR. At the same time, the ABC framework remains compatible with the principle of relevance as articulated in RT but relocates this principle within a dynamically structured organism–environment system in which relevance is enacted through graded modulation of affective, behavioral, and conceptual priors. The ABC framework thus extends standard RT by treating representational content as emergent from relevance-sensitive engagement rather than as the mechanism that makes such engagement possible.

5.1. ***Affective, Behavioral, and Cognitive Layers***

ABC shares a common philosophical ground with sensorimotor enactivism (Bruineberg 2017; Jaeger et al. 2024; Rietveld & Kiverstein 2014; Vervaeke et al. 2012), particularly the idea that meaning emerges through embodied engagement with the world rather than through internal representations (Kirchhoff & Kiverstein 2019). The A- and B-layers are dynamically primary, where the B-layer expresses these basic sensorimotor contingencies and the A-layer introduces affective valence and modulation, shaping salience, readiness, and directedness of action. The C-layer captures the wider conceptual, normative, and linguistic structures that stabilize long-horizon patterns of anticipation. However, the C-layer should not be understood as a storehouse of internal symbols, but as the temporally extended stabilization of normatively articulated practices emerging from affectively and behaviorally organized engagement. While enactivism emphasizes ongoing organism–environment coupling, the ABC framework adds a structured way of understanding how affective modulation and conceptual regulation co-govern the dynamics of that coupling.

ABC layers reflect depths of inference and temporal organization, explaining anticipation and error reduction through priming mechanisms. Different types of priming distribute across the layers:

1. A-priors correspond to affective priming, shaping motivational salience and readiness.

2. B-priors correspond to perceptual, semantic, and procedural priming, stabilizing habits and skills.

3. C-priors correspond to contextual and pragmatic priming, guiding slow, norm-sensitive inference.

Note that these priors are not stored propositions but embodied dispositions, action-shaped and affect-shaped tendencies within a dynamical repertoire.

The Theory of Mind (ToM) distinctions can also be reinterpreted through this lens. While the ToM taxonomy classifies *levels of mentalizing*, the ABC framework instead describes *levels of reorganization* required for an agent to maintain a grip on an unfolding situation. Low-depth episodes, dominated by A- and B-layer dynamics, align with hot ToM: fast, affectively modulated, embodied sense-making that leverages situated cues rather than detached reasoning. High-depth episodes, dominated by C-layer regulation, correspond to cold ToM: conceptually scaffolded reorientation, counterfactual sensitivity, and norm-guided deliberation.

Affect in the ABC architecture functions as a regulatory state that operates via *precision-weighted modulation*: each active state includes an implicit estimate of how much confidence to place in current predictions versus incoming sensory evidence or habitual behavioral tendencies (Yon & Frith 2021). When affective precision is high, intuitive patterns of engagement dominate; when affective precision is down-regulated, the system opens space for exploratory or reflective reorganization.

Within this framework, higher-order thought is not a representational meta-state but the expression of a shift toward C-layer re-coordination. The agent *reorganizes* its ongoing affective and behavioral tendencies so that previously suppressed or weakly weighted action-possibilities become salient enough to guide deliberative choice. Conscious deliberation thus emerges from the controlled redistribution of precision across layers: by loosening the grip of entrenched affective–behavioral tendencies and stabilizing wider temporal horizons of anticipation, a translator may bring alternative courses of action into experiential reach without constructing inner models of them.

The sarcasm example mentioned in 2.3.2 “Oh great, another meeting…” is not first encoded propositionally as “They are frustrated,” but is encountered through embodied cues, i.e., tone, prosody and posture, that draw the interlocutor into an expressive field. Understanding arises from attunement rather than from inference. Translation follows the same logic. Translators, like any engaged listener, resonate with the speaker’s expressive dynamics; understanding emerges from the interplay of sensorimotor attunement, affective resonance, and familiarity with cultural and linguistic affordances. As Sturm (2017, 420) observes, translators describe translation as “an act of supreme empathy,” despite its seemingly solitary nature. What appears as interaction with “just a text” is in fact a rich form of interpersonal and cultural engagement, enacted through embodied attunement, affective sensitivity, and the situated coordination of cognitive resources over time.

5.2. ***ABC Layers in Metaphor Translation***

Within the ABC framework, the translation of Chinese metaphorically used color terms offers a concrete demonstration of how different types of priming scaffold distinct translation strategies without invoking representational content. The following examples show how translators oscillate among A-, B-, and C-layer engagements depending on which dispositions are activated in real time.

Take *黑着脸* hēi zhe liǎn [lit. with a darkened face] as an example. The expression primes a bodily affective configuration of tightened facial muscles, suppressed tension, and a socially recognizable posture of displeasure. Depending on context and the translator’s situational engagement, this priming can activate different layers in the ABC taxonomy: perceptual coupling with visual darkness (B-layer), affective resonance with interpersonal tension (A-layer), or normative awareness of conventional emotional expression patterns (C-layer).

B-layer perceptual priming is often initially dominant. When translators first encounter this expression, the visual stimulus *黑* hēi [black] immediately activates perceptual associations like darkness, shadow and visual intensity, leading to a surface-proximal rendering that maintains the color metaphor: “He walked in *with a black face*.” This reflects sensorimotor coupling with

the visual-semantic structure of the source text rather than deeper affective or cultural engagement.

When C-layer procedural priming becomes activated, specifically, normative patterns concerning how emotional expressions are conventionally rendered across Chinese and English, the translator may shift toward a more interactionally attuned strategy such as "He walked in *looking stern*" or "*with a dark expression*." Here, translation arises from situational attunement to familiar interpersonal patterns rather than from manipulating abstract emotional concepts.

Under stronger A-layer affective priming, the translator becomes responsive to the affective contour enacted by the phrase (e.g., anger, pressure, interpersonal tension) and such embodied resonance may yield a rendering like "He walked in *with barely concealed anger*" or "*seething with displeasure*."

The example shows how the ABC framework can fluidly model shift between surface replication, normative alignment, and affective intensification within a single unfolding episode. Taken together, these variants illustrate how priming gradients evolve dynamically: a translation may begin with B-layer perceptual resonance, shift toward C-layer normative and procedural alignment, and intensify into A-layer affective attunement. The resulting translation strategies reflect dynamic patterns of embodied engagement shaped by the interaction of priming across ABC layers, stabilizing meaning into representational structures.

6. **Conclusion**

The preceding sections have advanced two interconnected claims. For one thing, standard RT treats nested mental representations as the primary engine of translation. We suggest, instead, to conceptualize translation as norm-sensitive realization of relevance through dynamically weighted structural alignment that later stabilizes into representational forms. For another, relevance itself can be understood within an enacted and embodied account of sense-making, structured through the ABC framework. Rather than treating representational content as explanatorily primary, we argue that representational stabilization should be understood as a higher-level achievement emerging from more fundamental affective, behavioral, and interactional dynamics. Standard RT, with its distinction between S-mode and I-mode translation,

already points in this direction, insofar as it acknowledges that translation is not merely the transfer of propositional content, but allows for different regimes of normative commitment.

6.1. ***S-Mode and I-Mode as Regimes of Engagement***

Under the representational assumptions in standard RT, S-mode translation appears as stimulus-oriented minimal reconstruction of intended thought, whereas I-mode translation involves inferential modeling of speaker intentions and audience beliefs. Our enactivist reinterpretation reframes this contrast.

S-mode can be understood as tight sensorimotor coupling with linguistic affordances, guided primarily by perceptual and structural priming. The translator's engagement remains close to the surface features of the source text, where interpretive resemblance emerges through stabilized patterns of embodied responsiveness rather than through explicit representational matching.

I-mode, in contrast, reflects deeper reorganization within culturally scaffolded normative space. Here, affective orientation, contextual sensitivity, and long-horizon anticipatory structures modulate translational decisions. This reorganization may involve reflective awareness and justification, but it need not be explained as recursive manipulation of nested mental representations. Instead, it can be described as a redistribution of precision across affective, behavioral, and conceptual layers within an evolving organism–environment system.

On this reading, interpretive resemblance is enacted alignment rather than representational mirroring. Representational stabilization may occur at later stages, particularly when translators articulate or justify their choices, but it is not assumed to be the generative core of translational sense-making.

6.2. *What the ABC Framework Contributes*

The ABC framework models translation as temporally layered, relevance-sensitive sense-making across affective (A), behavioral–sensorimotor (B), and cognitive–conceptual (C) dynamics. This layered account allows us to integrate phenomena that are sometimes treated as secondary within representational descriptions: affective modulation, priming gradients, habitual skill, and norm-sensitive stabilization. While representational models can and often do incorporate priming and automaticity, they typically describe these processes within a vocabulary of mental content and inferential operations. RR relocates the explanatory emphasis: rather than beginning with contentful representations that are subsequently modulated, it begins with precision-weighted organism–environment coupling.

Relevance is no longer understood primarily as optimization over pre-existing representations, but as a dynamic regulation of salience within a living agent embedded in a structured environment. Representational structures appear as sedimented outcomes of relevance realization, which unfolds as normatively stabilized patterns. The ABC framework makes it explicit that stabilized representational structures may emerge from the interaction between the three layers which can then be articulated and justified at the C-layer.

Translators frequently describe their work in terms of empathy, immersion, flow, and affective attunement. Standard RT often categorizes such experiences as recursive belief attribution (Sturm 2020; Gallai 2022). However, from an ABC perspective this is a miscategorization, since beliefs are commonly understood to emerge on the C-layer. The ABC framework, in contrast, incorporates these affective dimensions through precision-weighting and embodied responsiveness as constitutive features of translation on the A-layer, rather than as peripheral additions to inferential modeling.

6.3. *Synthesizing Translation Modes with ABC Layers*

This ABC framework allows us to map S-mode and I-mode translation practices onto specific configurations of layer activation and priming dominance:

1. S-mode translation operates primarily through B-layer perceptual and structural priming, maintaining tight coupling with surface linguistic features without engaging deeper affective or normative resources.
2. I-mode translation with cold ToM recruits B/C-layer coordination through procedural priming, drawing on entrenched translation routines and probabilistic expectations from experience.
3. I-mode translation with hot ToM activates A/C-layer coupling through affective priming, engaging culturally shaped, valence-based readiness to interpret tone, attitude, and stance.
4. I-mode translation with HOToM involves full-depth C-layer regulation through contextual and pragmatic priming, recruiting genre expectations, discourse norms, and meta-level priors about the communicative situation.

Table 1: S-mode translation, I-mode translation and ToM under an enactivist interpretation

Mode	Enactivist ABC Interpretation	ABC Layer(s)	Dominant Priming Type
S-mode (non-ToM)	Embodied reactivity to linguistic form; surface-level prediction	B-layer	Perceptual & semantic priming
I-mode (cold ToM)	Skillful, habitual sensorimotor attunement to communicative cues	B/C interface	Procedural priming
I-mode (hot ToM)	Affective co-regulation; emotion-guided sense-making	A/C coupling	Affective priming
I-mode (HOToM)	Deep participatory sense-making; metacoupling with communicative ecology	Full C-layer	Contextual & pragmatic priming

Table 1 maps S-mode and I-mode translation onto ToM categories in an enactivist interpretation. It offers an integrated account of affective modulation, priming dynamics, temporal unfolding, and norm-sensitive stabilization while preserving RT's central insight that communication is governed by relevance.

6.4. ***Implications for Cognitive Translation Studies***

In this paper, we suggest a view in which translation is seen as an enacted and relevance-guided coordination of (A) affective, (B) behavioral, and (C) cognitive dynamics, under temporally extended episodes of relevance realization. Translation process data such as keystrokes and gazes can provide empirical evidence for shifts between the ABC layers.

Future research in Cognitive Translation Studies could investigate how task complexity, affective valence, and normative expectations modulate cross-layer dynamics in real time.

References

Berger, Jacob, and Richard Brown. 2022. “Rosenthal’s Representationalism.” In *Qualitative Consciousness: Themes From the Philosophy of David Rosenthal*, edited by Weisberg Josh, 123–141. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108768085.010.

Bruineberg, Jelle. 2017. “Active Inference and the Primacy of the ‘I Can’.” In *Philosophy and Predictive Processing*, edited by Thomas K. Metzinger, and Wanja Wiese. Frankfurt am Main: MIND Group. doi: 10.155027/9783958573062.

Butterfill, Stephen A., and Ian A. Apperly. 2013. “How to Construct a Minimal Theory of Mind.” *Mind and Language*, 28(5): 606–637. https://doi.org/10.1111/mila.12036.

Carl, Michael, Takanori Mizowaki, Aishvarya Raj, Masaru Yamada, Devi Sri Bandaru, Yuxiang Wei, and Xinyue Ren. 2025. “From Representation to Enactment: The ABC Framework of the Translating Mind.” Pre-print, https://doi.org/10.48550/arXiv.2511.16811.

Carl, Michael. 2024. “An Active Inference Agent for Modeling Human Translation Processes.” *Entropy* 26 (8): 616. https://doi.org/10.3390/e26080616.

Carl, Michael. 2025a. “Representation and Resemblance in Translation: Scrutinizing Interpretive Language Use in Relevance Theory.” In *Applications of Relevance Theory to Translation and Interpreting*, edited by Alves Fabio, and Gallai Fabrizio, 41–62. New York: Routledge. https://doi.org/10.4324/9781003533245-4.

Carl, Michael. 2025b. “Temporal Dynamics of Emotion and Cognition in Human Translation: Integrating the Task Segment Framework and the HOF Taxonomy.” *Digital Studies in Language and Literature* 2 (2): 202–222. https://doi.org/10.1515/dsll-2025-0002.

De Jaegher, Hanne, and Ezequiel Di Paolo. 2007. “Participatory sense-making: An enactive approach to social cognition.” *Phenomenology and the Cognitive Sciences* 6(4): 485–507. https://doi.org/10.1007/s11097-007-9076-9.

Draperi, Manali, Ania Aïte, Mathieu Cassotti, Lorna Le Stanc, Olivier Houdé, and Grégoire Borst. 2022. “Development of cool and hot theory of mind and cool and hot inhibitory

control abilities from 3.5 to 6.5 years of age." *PLoS ONE* 17(1): e0262251. https://doi.org/10.1371/journal.pone.0262251.

Gallagher, Shaun. 2001. "The practice of mind: Theory, simulation, or interaction?" In *Between ourselves: Second-person issues in the study of consciousness*, edited by Thompson Evan, 83–108. Exeter: Imprint Academic.

Gallagher, Shaun. 2017. *Enactivist Interventions: Rethinking the Mind.* New York: Oxford University Press.

Gallai, Fabrizio. 2022. *Relevance Theory in Translation and Interpreting: A Cognitive-Pragmatic Approach*. 1st Edition. New York: Routledge. https://doi.org/10.4324/9781003183969.

Gładziejewski, Paweł. 2016. "Action guidance is not enough, representations need correspondence too: A plea for a two-factor theory of representation." *New Ideas in Psychology* 40: 13–25. https://doi.org/10.1016/j.newideapsych.2015.01.005.

Gutt, Ernst-August. 1991. *Translation and relevance: Cognition and context.* Oxford: Blackwell.

Gutt, Ernst-August. 2000. *Translation and relevance: Cognition and context.* 2nd Edition. Manchester: St. Jerome Publishing.

Gutt, Ernst-August. 2004. "Challenges of metarepresentation to translation competence." In *Translationskompetenz: Proceedings of LICTRA 2001: VII. Leipziger Internationale Konferenz zu Grundfragen der Translatologie*, edited by E. Fleischmann, P.A. Schmitt, & G. Wotjak, 77–89. Tübingen: Stauffenburg.

Gutt, Ernst-August. 2005. "On the Significance of the Cognitive Core of Translation." *The Translator* 11(1): 25–49.

Hirshkowitz, Amy, and M.D. Rutherford. 2021. "Longer looking to agent with false belief at 7 but not 6 months of age." *Infant and Child Development* 30 (5): e2263. https://doi.org/10.1002/icd.2263.

Hoza, Jack. 2016. *Interpreting in the Zone: How the Conscious and Unconscious function in Interpretation*. Washington, D.C.: Gallaudet University Press. https://doi.org/10.2307/j.ctv2rh29ff.

Hutto, Daniel D., and Erik Myin. 2012. *Radicalizing enactivism: Basic minds without content.* Cambridge: MIT Press.

Hutto, Daniel D., and Erik Myin. 2017. *Evolving enactivism: Basic minds meet content.* Cambridge: MIT Press.

Jaeger, Johannes, Anna Riedl, Alex Djedovic, John Vervaeke, and Denis Walsh. 2024. "Naturalizing relevance realization: why agency and cognition are fundamentally not computational." *Frontiers in Psychology* 15. https://doi.org/10.3389/fpsyg.2024.1362658.

Kirchhoff, M.D., and Kiverstein, Julian. 2019. *Extended Consciousness and Predictive Processing: A Third-Wave View*. London: Routledge. https://doi.org/10.4324/9781315150420.

Kiverstein, Julian, and Erik Rietveld. 2018. "Reconceiving representation-hungry cognition: an ecological-enactive proposal." *Adaptive behavior* 26 (4): 147–163. https://doi.org/10.1177/1059712318772778.

Kiverstein, Julian, and Erik Rietveld. 2021. "Scaling-up skilled intentionality to linguistic thought." *Synthese 198* (Suppl 1): 175–194. https://doi.org/10.1007/s11229-020-02540-3.

Lee, Jonny, and Daniel Calder. 2023. "The many problems with S-representation (and how to solve them)." *Philosophy and the Mind Sciences* 4. https://doi.org/10.33735/phimisci.2023.9758.

McCrea, Simon M. 2010. "Intuition, insight, and the right hemisphere: Emergence of higher sociocognitive functions." *Psychology Research and Behavior Management* 3: 1–39. https://doi.org/10.2147/PRBM.S7935.

Mercier, Hugo, and Dan Sperber. 2017. *The enigma of reason*. Cambridge: Harvard University Press. https://doi.org/10.4159/9780674977860.

Origgi, Gloria, and Dan Sperber. 2000. "Evolution, communication, and the proper function of language." In *Evolution and the Human Mind: Modularity, Language and Meta-Cognition*., edited by Peter Carruthers, and Andrew Chamberlain, 140–169. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511611926.008.

Premack, David, and Guy Woodruff. 1978. “Does the chimpanzee have a theory of mind?” *Behavioral and Brain Sciences* 1 (4): 515–526. https://doi.org/10.1017/s0140525x00076512.

Prinz, Jesse. (2004). “Which Emotions are Basic?” In *Emotion, Evolution, and Rationality*, edited by Dylan Evans, and Pierre Cruse, 69–88. Oxford: Oxford University Press.

Ramstead, Maxwell JD, Michael D Kirchhoff, and Karl J Friston. 2019. “A Tale of Two Densities: Active Inference Is Enactive Inference.” *Adaptive Behavior* 28 (4): 225–239. https://doi.org/10.1177/1059712319862774.

Rietveld, Erik, and Julian Kiverstein. 2014. “A Rich Landscape of Affordances.” *Ecological Psychology* 26 (4): 325–352. https://doi.org/10.1080/10407413.2014.958035.

Rietveld, Erik, Damiaan Denys, and Maarten Van Westen. 2018. “Ecological-Enactive Cognition as engaging with a field of relevant affordances: The Skilled Intentionality Framework (SIF).” In *The Oxford Handbook of 4E Cognition*, edited by Albert Newen, Leon De Bruin, and Shaun Gallagher, 41–70. New York: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198735410.013.3.

Robinson, Douglas. 2022. *Priming Translation: Cognitive, Affective, and Social Factors.* New York: Routledge. https://doi.org/10.4324/9781003134312.

Robinson, Douglas. 2023. *Questions for Translation Studies.* Philadelphia: Benjamins Translation Library. https://doi.org/10.1075/btl.162.

Rosenthal, David M. 2005. *Consciousness and Mind*. Oxford: Clarendon Press.

Schaeffer, Moritz, and Michael Carl. 2013. “Shared representations and the translation process: A recursive model.” *Translation and Interpreting Studies* 8 (2): 169–190. https://doi.org/10.1075/tis.8.2.03sch.

Sperber, Dan, and Deirdre Wilson. 1986. *Relevance: Communication and Cognition.* Oxford: Blackwell. 2nd edition published in 1995.

Sperber, Dan. 1997. “Intuitive and Reflective Beliefs.” *Mind and Language* 12 (1): 67–83.

Sturm, Annegret. 2017. “Metaminds: Using metarepresentation to model minds in translation.” In *Empirical modelling of translation and interpreting*, edited by Silvia Hansen-Schirra,

Oliver Czulo, and Sascha Hofman, 419–439. Berlin: Language Science Press. https://doi.org/10.5281/zenodo.1090990.

Sturm, Annegret. 2020. *Theory of Mind in Translation*. Berlin: Frank & Timme.

Szpak, Karina Sarto, Fabio Alves, Narhalia Biachini Esper, and Augusto Buchweiz. 2025. "From describing state of affairs to interpreting attributed thoughts: A relevant-theoretic interpretation of brainimaging data." In *Applications of Relevance Theory to Translation and Interpreting: Perspectives on Theory, Research and Practice*, edited by Fabio Alves and Fabrizio Gallai, 79–110. New York: Routledge. https://doi.org/10.4324/9781003533245-6.

Varela, Francisco J., Eleanor Rosch, and Evan Thompson. 1991. *The embodied mind: Cognitive science and human experience*. Cambridge: MIT Press. https://doi.org/10.7551/mitpress/6730.001.0001.

Vervaeke, John, Timothy P. Lillicrap, and Blake A. Richards. 2012. "Relevance Realization and the Emerging Framework in Cognitive Science." *Journal of Logic and Computation* 22 (1): 79–99. https://doi.org/10.1093/logcom/exp067.

Wilson, Deirdre. 2000. "Metarepresentation in Linguistic Communication." In *Metarepresentations: A Multidisciplinary Perspective*, edited by Dan Sperber, 411–448. New York: Oxford University Press. https://doi.org/10.1093/oso/9780195141146.003.0015.

Yon, Daniel, and Chris D. Frith. 2021. "Precision and the Bayesian Brain." *Current Biology* 31 (17): 1026–1032. https://doi.org/10.1016/j.cub.2021.07.044.

Authors' Bio

Dr. Michael Carl is a Distinguished Professor at Kent State University and Director of the Center for Research and Innovation in Translation and Translation Technology (CRITT). He has worked and published for more than 25 years in the fields of machine translation, computational linguistics, and translation process research. His work in the past decade was mainly centered around the conceptualization, analysis, and evaluation, as well as the empirically grounded modelling of the CRITT TPR-DB data.

Xinyue Ren is a PhD candidate from City University of Hong Kong. Her research focuses on metaphor translation, translation process research and cognitive linguistics. She is a member of the Association for Researching and Applying Metaphor (RaAM).