https://doi.org/10.21555/top.v740.3138

Skeletons in the Closet of Natural Language Semantics: Hidden Transcendental Assumptions of Donald Davidson's and David Lewis's Project

Los secretos ocultos de la semántica del lenguaje natural: presupuestos trascendentales del proyecto de Donald Davidson y David Lewis

Lucas Vollet
Federal University of Santa Catarina
Brazil
luvollet@gmail.com
https://orcid.org/0000-0002-6300-491X

Recibido: 22 - 04 - 2024.
Aceptado: 02 - 06 - 2024.
Publicado en línea: 10 - 11 - 2025.

Cómo citar este artículo: Vollet, L. (2026). Skeletons in the Closet of Natural Language Semantics: Hidden Transcendental Assumptions of Donald Davidson's and David Lewis's Project. *Tópicos, Revista de Filosofía*, 74, 231-260. https://doi.org/10.21555/top.v740.3138

Abstract

This article provides a critique of Davidson's and Lewis's dependence on truth conditions in natural language semantics, connecting their shortcomings to broader epistemological challenges. Despite their formalist aspirations and claims of epistemic neutrality—centered on a semantic mapping purportedly indifferent to pure categories or specific judgment standards—we argue that both frameworks ultimately rely on what Gross (2012) terms the "Principle of Exhaustion," which limits the scope of meaning to publicly available semantic facts, collapsing semantic agreement into a surface-level compatibility devoid of deeper epistemic divergence and normative content. We argue that both Davidson's extensionalist theory and Lewis's intensionalist theory rely on underlying assumptions about what counts as rational, judgeable, or meaningful within a given framework of truth standards, ultimately failing to avoid inclusion in a transcendental problematic. Both programs fail to remain transcendentally neutral or indifferent regarding the epistemic standards that structure the limits of judgment (or any method for determining truth), a failure that manifests either in (1) their reduction to a merely formal paradigm incapable of sustaining the transition from a theory of *contentless agreement* to a theory of meaning, or (2) their slipping into an unreflective demarcation of meaning, shaped dogmatically and blindly by the prevailing paradigms of method and judgment of a given historical period.

Keywords: natural language semantics; meaning theories; transcendental assumptions; truth; truth conditions; epistemic agreement; Principle of Exhaustion; Donald Davidson; David Lewis; Immanuel Kant.

Resumen

Este artículo critica la dependencia de Davidson y Lewis a las condiciones de verdad en la semántica del lenguaje natural relacionando sus deficiencias con retos epistemológicos más amplios. A pesar de sus aspiraciones formalistas y sus afirmaciones de neutralidad epistémica —centradas en un mapeo semántico supuestamente indiferente a las categorías puras o a los estándares de juicio específicos—, sostengo que ambos marcos se basan en última instancia en lo que Gross (2012) denomina el "Principio de Agotamiento", que limita el alcance del significado a los hechos semánticos disponibles públicamente, reduciendo el acuerdo semántico a una compatibilidad superficial desprovista de divergencias epistémicas más profundas y de contenido normativo. Sostengo que tanto la teoría extensionalista de Davidson como la teoría intensionalista de Lewis se basan en supuestos subyacentes sobre lo que se considera racional, juzgable o significativo dentro de un marco de estándares de verdad, sin poder evitar, en última instancia, su inclusión en una problemática trascendental. Ambos programas fracasan en su intento de permanecer trascendentalmente neutrales o indiferentes con respecto a los estándares epistémicos que estructuran los límites del juicio (o cualquier método para determinar la verdad), un fracaso que se manifiesta bien en (1) su reducción a un paradigma meramente formal incapaz de sostener la transición de una teoría del acuerdo sin contenido a una teoría del significado, bien en (2) el deslizamiento hacia una demarcación irreflexiva del significado, moldeada dogmática y ciegamente por los paradigmas predominantes de método y juicio de un periodo histórico dado.

Palabras clave: semántica del lenguaje natural; teorías del significado; supuestos trascendentales; verdad; condiciones de verdad; acuerdo epistémico; principio de agotamiento; Donald Davidson; David Lewis; Immanuel Kant.

1. Introduction

Donald Davidson's theory, while formally inspired by Tarski's conception of truth, represents a significant departure in its application to natural language. Tarski himself approached natural languages with caution. For Tarski, these features introduced complexities that resisted straightforward formalization, leading him to focus primarily on formalized languages where such issues could be more readily controlled.

Davidson, however, a pioneer among a group of theorists equally committed to the ambitious project of elucidating meaning through formal structures, viewed these complexities not as obstacles but as essential features of language itself. His use of T-sentences—statements of the form, "'p' is true if and only if p"—was not merely an adaptation of Tarski's formal precision but an innovative attempt to reconcile it with the interpretive challenges posed by natural language. For Davidson, these sentences served as conceptual bridges, anchoring the systematicity of truth-theory to the fluid and context-sensitive nature of human communication. In his view, the apparent messiness of natural language was not a defect to be circumvented but a phenomenon to be accounted for, demonstrating the interplay between structure and interpretation that lies at the heart of linguistic understanding.

Davidson's bold endeavor was encouraged by what was called by Steven Gross (2012) the "Principle of Exhaustion", which holds that all semantic facts about meaning are publicly accessible through a speaker's observable behavior and the context of their utterances, leaving no room for hidden or inaccessible aspects of meaning. The novelty of this principle lies in including an epistemic condition among the requirements Davidson sets for a theory of meaning: "The evidence available to a radical interpreter, given the constraints to which radical interpretation is subject, exhausts the relevant semantic facts, in the sense both of determining them and rendering them epistemically determinable" (Gross, 2012, p. 228). In this article, we assume that this principle operates implicitly, reinforcing Davidson's confidence in the irrelevance of the epistemic dimension of a theory of meaning. That is, Davidson appears to treat epistemic considerations as fully subsumed under semantic facts, effectively bypassing any need to critically address how meaning depends on deeper epistemic structures. This principle

assumes that meaning is fully manifest in linguistic usage, thus enabling a systematic truth-theoretic approach to language. Semantic theories based on this principle aim to ensure consistency and coherence in interpretation, aligning the formal structure of truth conditions with observable linguistic practices.

While other authors that engaged in the project of natural language semantics have developed approaches that relax Davidson's anti-modal and extensional constraints, their theories nevertheless reflect an implicit reliance on the Principle of Semantic Exhaustion. David Lewis's work offers a noteworthy comparison. Unlike Davidson, who systematically avoids modal and intensional contexts, Lewis explicitly embraces these through his possible-worlds framework, accommodating counterfactuals and contexts of intensionality. His emphasis on compositional structures and the grounding of meaning in conventions showcases a broader ambition for semantic generality. As Lewis states, "Semantics with no treatment of truth conditions is not semantics" (1970, p. 19), echoing Davidson's central commitment to truth conditions. This approach builds on a rich tradition of referential semantics influenced by Frege (1892), Tarski (1956), Carnap (1947 & 1963), Kripke (1963), Kaplan (1964), Montague (1960, 1968 & 1971), and Scott (1970), integrating their insights. Lewis's framework expands on the methodologies permitted for a truth-based semantics, including meaning defined as functions from indices to truth values, aligning linguistic expressions systematically with extensions across possible scenarios: "Generalizing over the time and assignment coordinates and letting the others (including world) be determined by context, we define *eternal truth* of a sentence meaning on an occasion; generalizing over the assignment coordinate and letting all the rest be determined by context, we define simply truth on an occasion; and so on" (Lewis, 1970, p. 34). Despite his broader semantic scope, Lewis, like Davidson, operates within a framework where meaning risks being reduced to mere compatibility with truth values, even though his methodologies are more expansive. In the following sections, we will argue in greater detail how this aligns Lewis with the Principle of Exhaustion, effectively placing him on similar theoretical footing.

However, this commitment to semantic exhaustion reveals significant limitations, particularly through what might be termed the *Problem of Contentless Agreement*. The Principle of Exhaustion assumes that all semantic facts are publicly observable, accessible through

behavioral evidence, and learnable as patterns of assent and dissent. While David Lewis explicitly rejects conflating formal semantics with psychology or sociology, his framework ultimately aligns with a view where the structure of a grammar is exhaustively learnable as patterns that fully account for semantic phenomena. As he states,

> I distinguish two topics: first, the description of possible languages or grammars as abstract semantic systems whereby symbols are associated with aspects of the world; and second, the description of the psychological and sociological facts whereby a particular one of these abstract semantic systems is the one used by a person or population. Only confusion comes of mixing these two topics (Lewis, 1970, p. 19).

The frameworks of both Davidson and Lewis make the structure of grammar learnable as patterns that fully encompass semantic content, as they share the premise that any language—formal or natural—remains semantically predictable as long as it avoids unlearnable elements or unnecessary structural complexity beyond what semantic facts require. This premise ensures that semantics is tied to observable patterns of use, aligning with their commitment to grounding language in learnable systems.

While this framework has pragmatic appeal, the reliance on exhaustively accessible formal structures risks reducing meaning to learnable patterns without addressing deeper, irreducible aspects of semantic understanding. It strips meaning of its richer epistemic and normative dimensions—those that provide substance to divergences between conceptual schemes, webs of belief, and other elements that ground divergence among interpretive frameworks. For instance, the contrast between the paradigm of Lamarckian inheritance and Darwinian evolution in biology highlights how different scientific frameworks influence meaning. Lamarck's theory posited that acquired traits could be passed to offspring, grounded in an assumption of direct environmental influence on heredity. In contrast, Darwinian evolution, supported by Mendelian genetics, emphasized natural selection and random genetic variation as the drivers of inherited traits. These paradigmatic shifts reflect not just factual disagreements but deeper epistemic and normative differences in how biological phenomena are understood and explained—dimensions that cannot be captured solely

by reducing meaning to compatibility with observable patterns or truth values.

In Davidson's work, such differences become invisible, which he celebrates as a triumph over the turbulent waters of incommensurability and conceptual relativism. However, this victory comes at a cost: by dismissing these differences as irrelevant or incoherent, the theory risks flattening the complexity of meaning into a surface-level agreement, devoid of the depth that such divergences might otherwise reveal. The result is that the agreement we rely on when interpreting and communicating meaning is ultimately hollow: it is reduced to a narrowly defined margin of truth where anything compatible becomes indistinguishable. Both Davidson and Lewis, despite their differences, risk collapsing meaning into formal compatibility with truth conditions, where the richness of semantic nuance is flattened into patterns of agreement constrained by learnable rules.

This creates what can be termed the self-insufficiency of semantics—a failure to serve as an exhaustive representation of the phenomenon of meaning as it develops in parallel with the forms of cognitive access made possible by prevailing epistemic paradigms – i.e., a framework or system of knowledge that shapes how we understand, interpret, and acquire knowledge about the world. Epistemic paradigms encompass the dominant theories, methodologies, assumptions, and practices that shape what is regarded as valid or credible knowledge within a specific historical or intellectual context. The formal system, which aspires to be self-contained and exhaustive, ultimately depends on unacknowledged epistemic paradigms to address its shortcomings. These paradigms act as an implicit scaffolding guiding decisions and interpretations beyond the scope of the formal structure. These paradigms, in turn, determine the criteria for selecting the most coherent and effective "webs of belief" in a given historical moment. The insufficiency of formal frameworks in semantics becomes evident when they encounter situations of indecision, uncertainty, or ambiguity. Even when all semantic facts are exhaustively accounted for, these frameworks often fail to provide clear or definitive answers through their internal rules and mechanisms alone.

As we will argue, natural language semantics frameworks, such as those proposed by Davidson and Lewis, share a fundamental limitation: they reduce meaning to a mere contentless agreement rooted in formal regularities, overlooking the epistemic framework that enables communication and interpretation. This article seeks to

suggest that natural language semantics theories, despite their claims to neutrality and objectivity, cannot escape a dependency on epistemic and normative preconditions that are inherently transcendental in nature. The failure to address these transcendental assumptions manifests as specific symptoms: (1) a reduction to a purely formal paradigm that fails to bridge the gap between contentless agreement and a comprehensive theory of meaning, or (2) an uncritical restriction of meaning, shaped rigidly by the dominant methods and judgments of a particular historical moment.

2. T-schemes, meaning-reductionism and its pragmatic interpretation

Davidson's theory of interpretation and the semantics of natural language arises as the most expansive program of philosophical exploration of Tarski's semantic theory. The program commences with a reductionist perspective, which suggests that the notion of *meaning* can be elucidated by referring to a less complex concept, namely *truth*. This project is entirely at odds with the expectations of Alfred Tarski himself, who, as early as 1944, expressed concern that natural language was too imprecise to align epistemic notions of proof with an extensionally delimited concept of truth. Tarski recognized the limitations of applying his formally defined truth concept to natural languages. In this article, we reinterpret the perceived limitations of natural language as, in fact, its strengths—rooted in its inherent reliance on richer epistemic frameworks that encompass context, intention, and shared cultural norms. This demonstrates that Davidson's project starts in defiance of a critical point, which, as we will see by the end, is one of the reasons for its ultimate failure.

In fact, moving from a purely formal notion of truth to a robust theory of meaning—one that accounts for natural language events of meaning—involves significant conceptual leaps. For Davidson, the idea is not merely to define truth in isolation but to show how truth conditions can illuminate the way language conveys meaning, or, in his words, "how we can go from truth to something like meaning" (Davidson, 2001, p. 74). Reductionism in the philosophy of language often involves simplifying the complex, layered nature of "meaning" by reducing it to a basic, structural framework. This framework seeks to explain meaning not through its rich inferential connections or semantic nuances, but instead

by identifying bare, factual conditions under which a given statement or proposition (let's say "p") would be considered true. The encoding of a recursive enumerative procedure for generating well-formed sentences and consistent patterns of interpretation can be utilized as predictable data for a theory of meaning: "someone who knows a theory of truth for a language L would have enough information to interpret what speaker of L says. [...] [T]he empirical and formal constraints on a theory of truth sufficiently limit the range of the acceptable theories" (Davidson, 2001 , p. 224). As put forth by Rorty: "We emphasize a point made by Davidson himself: that a Tarskian truth definition is an empirical theory, designed to find an underlying order behind a lot of confusing uses" (1999, p. 65). So, unless the language is absolutely incomprehensible to the point of missing any type of compositional structure, "one is in position to derive a T-sentence, 's is true iff p' (ignoring context sensitivity for the sake of simplifying exposition) for any object language sentence s, which draws only on the content of the axioms" (Lepore & Ludwig, 2012, p. 3). As Lepore and Ludwig (2013, p. 178) note, "[a] plausible constraint on a constructive theory is that it entails, from axioms about semantical primitives, all true sentences of the form: *s means that p*".

Building on Tarski's insights, Davidson emphasizes the pragmatic and compositional utility of the T-scheme within his truth-conditional semantics. The T-scheme has been incredibly successful in providing a framework that forms the foundation of our comprehension regarding the role of "p" in a theory of language composition. However, this understanding is less about a substantial view of the meaning of "is true" and more about the expressive role that this predicate plays in linguistic practice. By grasping this framework, we gain insight into the situations where asserting "p" does not involve presuppositions that undermine the intended objective. As a result, we can determine the circumstances in which "p" can be effectively communicated and understood by anyone familiar with the language. If individuals who tend to believe that "p" is true in situations where it is not were the ones endorsing these statements in a language, the processes of communication and correcting misinterpretations would break down. Davidson acknowledges this fact, and therefore he states that we should assume just the opposite: "Charity is forced upon us; whether we like it or not, if we want to understand others, we must take it to be true in most matters" (Davidson, 2001, p. 197).

Furthermore, Davidson's emphasis on the practical utility of truth aligns with his broader philosophical project of radical interpretation. According to Davidson (1973), understanding a speaker's language requires attributing beliefs, desires, and intentions in a way that makes their utterances intelligible. This process presupposes that speakers aim to assert what they take to be true, even in cases where their reasons for belief might be opaque or contextually contingent.

We can now examine how Davidson proposes making the significant leap from truth to something as ambitious and complex as meaning. The perception of meaning as hermetic and scientifically elusive often arises from its grounding in shared paradigms of agreement within cultural and epistemic practices. Davidson circumvents the need for such presuppositions by adopting an alternative foundation: what Gross (as discussed in the introduction) terms the Principle of Exhaustion. This principle asserts that all relevant semantic facts are fully captured by observable language patterns and contextual factors, serving as the cornerstone for Davidson's theories of radical interpretation, triangulation, and his rejection of conceptual incommensurability. Davidson acknowledges that different truth-theories may yield incompatible truth conditions for the same sentence, leading to a degree of indeterminacy (Davidson, 1973). However, the Principle of Exhaustion helps address this issue by stipulating that if there is nothing beyond semantic facts, there can be nothing "in-between" or implicit in interpretation that could introduce further ambiguity, provided that the truth-theories coincide extensionally. This allows Davidson to contain indeterminacy within the realm of competing truth conditions, avoiding any deeper epistemic or interpretive disputes.

While Davidson concedes that meaning can be underdetermined by evidence—aligning with Quine's (1969) indeterminacy thesis—he reframes this underdetermination as a constructive feature rather than a skeptical problem. This reframing imposes a practical limit on the range of interpretations, ensuring that the framework remains functional for effective communication. The principle thus serves to confine semantic disagreement to observable regularities, preventing it from escalating into more profound epistemic challenges. By doing so, Davidson effectively neutralizes potential criticisms of his approach, safeguarding the coherence and applicability of radical interpretation as long as extensional agreement between truth-theories is maintained.

The leap from truth standards to meaning, it seems, is not made by some mysterious alchemical process, but through the methodical application of truth-conditional semantics. "Truth" appears as a component for analyzing that element of the representation of propositions that may be *learned as meaning*, a representation of the proposition's unchanging content—or that which cannot be reversed once the proposition is asserted.

3. Lewis and Davidson: meaning exhausted by truth standards

David Lewis's work on natural language semantics (1969 & 1975) is foundational in formalizing how meaning arises from the compositional structure of language. By developing a framework rooted in possible-worlds semantics, Lewis demonstrated how the meaning of complex expressions can be systematically determined by the meanings of their constituent parts and the way they are combined. His theory of conventions further grounds this semantic compositionalism in a naturalistic understanding of social behavior, where truthfulness and trust emerge as stabilizing forces in communication. Despite its metaphysical underpinnings, Lewis's approach shares with Davidson's radical interpretation the ambition to create a unified account of meaning grounded in observable and systematic principles.

While Davidson's project hinges on the principle of charity and the necessity of shared rationality for radical interpretation, David Lewis offers a complementary view that emphasizes the social dimension of linguistic practice. In his seminal work *Convention* (1969) and later reflections on metasemantics (Lewis, 1975), Lewis argues that meaning arises not from internalist mental structures but from a social contract among language users, grounded in mutual expectations of truthfulness and trust: "Grammar Γ is used by P if and only if Γ is a best grammar for a language L that is used by P in virtue of a convention in P of truthfulness and trust in L" (Lewis, 1975, p. 177). Here, Lewis shifts the focus from individual intentionality to communal practices that stabilize linguistic norms. For Lewis, intentionality itself is not a private, mental phenomenon but an interpretative sign that emerges from the optimal rationalization of an organism's actions. This pragmatic view resonates with interpretivist metasemantics, where the semantic role of words is determined by their usage within a linguistic community.

This conception aligns with Lewis's metaphysical stance on properties and universals, where "properties" are not seen as inherent

or metaphysically independent structures but as abstractions emerging from patterns of resemblance among particulars. In "New Work for a Theory of Universals", Lewis (1999) argues that universals should be understood as "natural properties" that supervene on the arrangement of particular things in the world. Similarly, in his account of linguistic conventions, meaning is depicted as emerging from patterns of communicative behavior. In both frameworks, what we consider "real" or "meaningful" is grounded in the observable regularities that structure our experience.

However, Lewis's view goes beyond describing mere factual regularities. For meaning to operate on a deeper, more explanatory level, these patterns must respond to predictions governed by laws, rather than being confined to surface-level regularities. In this sense, the explanatory power of meaning, much like the natural properties of universals, depends on its ability to articulate connections within a systematic framework of rules and principles. As Lewis himself notes, "we should be prepared to find that in a more adequate method, meanings may turn out to be complicated, infinite entities built up out of elements belonging to various ontological categories" (1970, p. 34).

Lewis's critique of *property skepticism* addresses a key question that both Quine and Davidson[1] struggled to answer: what is the semantic role of abstract terms like "properties," if they cannot be reduced to particular objects or truth-functional terms? In contrast to both Quine and Davidson, David Lewis argued that rejecting properties entirely impoverishes our understanding of linguistic competence. In his critique of Quine's desert landscapes, Lewis contended that abstract constructs like properties play an essential semantic role: "What is the semantic role of these words [properties]? If we are to do compositional semantics in the way that is best developed, we need entities to assign as semantic values to these words, entities that will encode their semantic roles"

[1] Davidson's elimination of intensional entities can be seen as a radical commitment to clarity and empirical grounding—an extension of the scientific attitude toward identifying patterns in the world rather than getting lost in speculative metaphysics. For Davidson, intensional entities such as properties, universals, and meanings as abstract objects belong to the realm of ungrounded theoretical constructs that create more confusion than insight. His approach to semantics, influenced by Quine's extensionalism and Tarski's formal definition of truth, is essentially pattern-seeking without metaphysical baggage.

(1999, p. 212). In Lewis's view, these abstract terms are indispensable for explaining how language users navigate a shared world of natural kinds and stable resemblances. His account, grounded in Humean supervenience, emphasizes that properties are nothing more than patterns of similarity that emerge from observable regularities.

Properties, for Lewis, are not mysterious intensional entities but rather natural classes that emerge from patterns of resemblance among particulars. These natural properties function as interpretative tools that help us categorize the world and stabilize linguistic conventions. Lewis offers a model of naturalism that is compatible with recognizing the role of conventions, normativity, and modal facts in linguistic practice. His use of possible world semantics further reinforces this point, as it provides a framework in which meaning depends not only on actual truth conditions but also on the broader range of possible scenarios in which a statement could be true, highlighting the normative dimension of interpreting meaning across varying contexts and counterfactual situations.

This assumption supports the pragmatic view that, if interpretable at all, beliefs must aim for truth, even in hyper-speculative or problematic settings. For example, in intensional contexts, where the truth of a statement depends on how its components are interpreted rather than their sheer extension (e.g., "John believes the morning star is bright," which hinges on John's belief system rather than objective facts), the grounds for asserting can be unstable. Similarly, in epistemically contested assertions, such as scientific debates involving competing paradigms (e.g., quantum mechanics versus classical mechanics in specific domains), the grounds for assertion evolve as evidence and interpretation shift. Despite these challenges, the theory presumes that beliefs, even in such unstable contexts, aim for predictive coherence and truth, revealing both its practical appeal and its limitations in addressing the complexities of speculative domains. The theory of truth's only non-negotiable assumption is that non-defeatist beliefs behave predictably, maintaining stability and coherence under reasonable conditions. Truth-conditional theories are pragmatically appropriate because they avoid postulating dogmatic or hermeneutic rules for judging or overinterpreting the content of propositional attitudes and beliefs.

The attempt to generalize the conditions of interpretation by unifying them under a single standard of truth represents, in many ways, an effort to bypass the challenges posed by distinct epistemic standards or

underlying presuppositions that complicate the unified representation of "meaning." This strategy, already present in the foundational work of Tarski's formal semantics (not for natural languages), seeks to establish an ostensibly neutral and universal framework for truth, thereby sidelining the sociohistorical and normative contexts that influence interpretation.

In this article, we aim to challenge the assumption that it is possible to completely abstract from these conditions. We argue that any such abstraction necessarily relies on implicit normative and epistemic frameworks, which themselves shape the processes of interpretation and meaning-making, and furthermore involve a degree of reflective commitment in the form of a transcendental theory (even if this remains hidden from the perspective of the technical practitioner).

4. Semantic exhaustion and the theory of radical interpretation: Davidson's critique of conceptual scheme dualism

Davidson's theory of radical interpretation emphasizes the conditions necessary for meaningful belief attribution, interpretation, and communication. Central to this framework is the presupposition that belief systems must be predominantly true to enable linguistic coherence, intersubjective communication, and effective engagement with the world.

Richard Rorty highlights this connection by situating Davidson's work within pragmatism. According to Rorty, Davidson's views, particularly his claim that "most of our beliefs must be true" and his rejection of a strict distinction between knowing a language and navigating the world, resonate more with Wittgenstein (2009) than with the systematic theories of Carnap or Tarski. Rorty explains:

> […] a lot of the doctrines for which Davidson became famous in the 1980s are more akin to those of the later Wittgenstein than to any views held by Carnap or Tarski. Consider his claims that most of our beliefs must be true, and that there is no distinction between knowing a language and knowing our way around the world. You can grasp his arguments for these claims even if you have no interest whatever in what Dummett called "a systematic theory of meaning"—the sort of theory which the later Wittgenstein thought implausible and

> unnecessary, but of which a Tarskian truth-theory for a natural language is paradigmatic (Rorty, 1999, p. 50).

Indeed, setting standards that enable us to align our own standards of truth with those of other community members is no different from setting standards to learn a language. Richard Rorty observed this point insightfully: "The systematic relations between linguistic expressions which are captured by the recursive character of a Tarskian truth definition are not different from the relations of being-frequently-inferred-from of which the radical interpreter, hoping to construct such a definition, must keep track" (Rorty, 1999, p. 68). Within the framework of radical interpretation theory, the parameters for interpreting "p" are idealized until any approach—arbitrary or not—for determining whether "p" is true is required to meet the standard by which "p" can be interpreted as true, that is, in accordance with the T-scheme's platitude. As Rattan (2005) clarifies:

> Davidson (1973) distinguishes his project of radical interpretation from Quine's radical translation; whereas a translation manual relates sentences to sentences, an interpretation relates sentences (or their parts) to worldly objects and events. Interpretation that culminates in instances of the T-schema as theorems of a meaning theory does just this: the use of the sentence on the right-hand side reports how things are just in case the sentence is true (Rattan, 2005, p. 54).

Because the method of radical interpretation excludes in advance any method of interpretation that places "p" on more onerous or substantial grounds than correspondence with p, it follows that learning to assert "p" in non-paradoxical situations, which is a prerequisite to understanding "p" in a language, will eventually converge with learning to assert "p" in the precise conditions in which to believe "that p" has a high probability of being true. As a result, we are simultaneously learning the most tranquil or prudent conditions for the statement of "p," those that would only turn "p" into a falsehood in the event of a drastic modification of the language's terms, or situations of extraordinary rupture.

Because Davidson's project operates with these presuppositions in place, he was able to bypass Tarski's concerns by using a truth theory constructed in formal terms as the foundation for a theory of meaning—

one that seeks to explain how meaning and agreement are understood in natural language. That is, Davidson seeks to explain all the epistemic nuances tied to the standards of agreement that a community relies on to establish a robust basis for meaning—ensuring that their judgments do not diverge so radically as to exceed the boundaries of a shared conceptual scheme—through a Tarskian framework, despite Tarski's own doubts about such an endeavor.

This transformation relies heavily on a crucial epistemological feature that Davidson adopts: the Principle of Exhaustion, as discussed by Steven Gross (2012). The Principle of Exhaustion (E) asserts that all relevant semantic facts about meaning are publicly available to a radical interpreter through observable behavior and circumstances. In other words, everything that is required to determine what a speaker means is already present in their actions and the context of their utterances, leaving no room for hidden, inaccessible aspects of meaning.

Thanks to the Principle of Semantic Exhaustion, Davidson seems to overlook the gap between "assertability" and "truth" at the ontological level, assuming that this approach is sufficient. Thanks to the Principle of Semantic Exhaustion, even if we acknowledge the infinite gap between "assertability" and "truth" at the ontological level, we might still assume that there are no cases of assertion that fall short of the project of reaching the truth—unless the speaker is irrational or we cannot extend charity to them for other reasons. Our finite nature prevents a perfect alignment between being justified (in assertion) and being true. However, the Principle of Exhaustion dismisses this as irrelevant, as it effectively decrees that we cannot separate our aim for truth from the assertive parameters used by anyone employing a non-defeatist (non-paradoxical) strategy to justify their beliefs. Thus, learning the meaning of p in a language L becomes indistinguishable from learning how to judge p in accordance with an acceptable level of public consensus and epistemic agreement.

But there is more to this self-confidence encouraged by the principle. For Davidson, the very act of successful interpretation demonstrates that divergent conceptual schemes are unnecessary. As he famously stated: "If we succeed in interpreting someone else then we have shown there is no need to speak of two conceptual schemes, while if we fail there is no ground for speaking of two" (Davidson, 1973, p. 20). Davidson's approach, particularly his critique of the "conceptual scheme-content dualism," dismisses the epistemic significance of differences shaped

by cultural, historical, or philosophical perspectives. All these forms of self-awareness regarding the foundations of agreement are rendered irrelevant at the epistemic level. Davidson even considers them to be misunderstandings, irrelevant confusions, or attempts to give substance to empty dualisms between conceptual schemes and empirical content—a "dogma of empiricism, the third dogma" (1973, p. 11):

> […] no substantive conclusions about similarities and dissimilarities between various ways of carving up the world follow from Davidson's reasoning. The real point of "On the Very Idea of Conceptual Scheme" is simpy that, if we are extensionalists about meaning, we cannot imagine what it is for a language to have an insight (Ramberg, 1989, pp. 120-121).

These differences—such as variations in concepts, norms, and practices that make certain interpretations natural or valid within one framework but alien or incoherent within another—are effectively rendered invisible in Davidson's theory.

5. The self-insufficiency of Semantics

We saw above that the principle of semantic exhaustion allows Davidson to adopt an overly confident stance, as it leads him to believe that his approach is sufficient—even when confronted with the ontological gap between "assertability" and "truth." While it is clear that our finite nature prevents the perfect alignment of being justified in an assertion and being true, the principle of exhaustion treats this discrepancy as irrelevant. It assumes that our pursuit of truth is inseparable from the assertive parameters employed by anyone using a non-defeatist strategy to justify their beliefs. Consequently, the process of learning the meaning of *p* in a language *L* collapses into the process of learning how to judge *p* based on a level of public consensus and epistemic agreement. This level is deemed sufficient for practical purposes and for a semantic account of natural language, which—unlike formal languages—is sensitive to epistemic standards of confirmation and verification that are not always formalizable. It is therefore unnecessary for speakers to learn anything beyond this: no specific scientific paradigm, no regional ontology, no set of Kantian categories, no standards of judgment. All of this is superfluous.

However, this approach reveals a significant limitation. The Principle of Exhaustion creates a contentless agreement: disputes about meaning are either dissolved into formal equivalence or dismissed as irrelevant because the framework lacks the resources to address them. This formal vacuum would not pose an issue if it were adequate to resolve all questions of meaning. Unfortunately, it falls short. Meaning, as experienced in real-world communication, encompasses far more than formal mechanisms; it is deeply shaped by shared epistemic norms.

When formal frameworks, even when all semantic facts are exhaustively accounted for, encounter situations of indecision, uncertainty, or ambiguity—where their rules and mechanisms fail to yield clear or definitive answers—they inevitably rely on broader, external standards for resolution. Consider a statement like "John holds that the morning star shines brightly." Here, the truth of the claim depends on John's personal belief system rather than solely on objective realities, rendering the basis for assertion inherently fluid. Likewise, in contexts of scientific disagreement—such as disputes between quantum mechanics and classical mechanics in particular fields—the foundation for asserting truth shifts over time as new evidence emerges and interpretations adapt to evolving paradigms. These epistemic standards are not inherent to the formal framework itself but are shaped by the dominant practices, assumptions, and norms of the specific historical and cultural context in which the framework operates. In other words, when the formal system cannot independently address a problem, it defaults to widely accepted ways of reasoning, judgment, and validation that reflect the prevailing epistemic paradigms of the time. These paradigms act as an unspoken foundation, guiding decisions and interpretations even though they lie outside the formal structure of the framework.

This insufficiency justifies the position of semantic skeptics who accepted Tarski's theory but rejected Davidson's program of deriving from it a comprehensive theory of meaning for natural language. Tarski himself would qualify as one of those skeptics, with his caution that he would not trust transitioning from his theory to a theory of communication of meanings in natural language. We can say that this is because natural language responds to varied epistemic standards, such as understanding an isolated word in relation to how it fits into a web of beliefs and how that web contrasts with others, rather than adhering solely to the idealized standards formalized in a T-schema.

6. The failure of transcendental neutrality in Davidson's and Lewis's semantic theories

In post-Kantian usage, the term "transcendental" refers to a reflection on the conditions of possibility for cognition and knowledge. In his words: "I call all cognition transcendental that is occupied not so much with objects but rather with our mode of cognition of objects insofar as this is to be possible a priori. A system of such concepts would be called transcendental philosophy" (*KrV* A12/B25). For Kant, "meaning" is intrinsically tied to the conditions that allow for the intuitive application of pure or empirical concepts, giving content to cognition. For him, since general logic fails to consider this content, and empirical investigation can only determine it *a posteriori*, a transcendental logic is necessary. This logic "has to do merely with the laws of the understanding and reason, but solely insofar as they are related to objects *a priori* and not, as in the case of general logic, to empirical as well as pure cognitions of reason without distinction" (A57/B82). Thus, the conditions of cognition are simultaneously the conditions of meaning. These conditions, according to Kant, are rooted in *a priori* pure forms or inherent frameworks that underpin paradigmatic principles of judgment, i.e., the principles involved in our distinction of truth: "although general logic can give no precepts to the power of judgment, things are quite different with transcendental logic, so that it even seems that the latter has as its proper business to correct and secure the power of judgment in the use of the pure understanding through determinate rules" (A135/B174). Davidson's reliance on Tarski's semantic theory seeks to establish a truth-conditional model of meaning—focused on observable and extensional criteria—while Lewis employs conventions and possible worlds to systematically account for linguistic patterns. Both thinkers emphasize methodological neutrality, avoiding explicit commitments to specific epistemic categories, standards of judgment, or, in Davidson's case, metaphysical assumptions. This gives us sufficient grounds to question whether these theories—Davidson's and Lewis's—exhibit any transcendental character or inclination.

Lewis's account explicitly rejects the transcendental framework by embracing two theses. For Lewis, conventions are understood as *a posteriori* social facts, while properties are framed as higher-order

realities that supervene on empirical arrangements. This naturalistic stance departs significantly from Kantian transcendentalism, which posits that the conditions of knowledge and meaning are grounded not in empirical contingencies but in the *a priori* structures of human cognition.

Furthermore, Lewis's endorsement of the metaphysical reality of universals directly contradicts a central tenet of transcendentalism: the incognoscibility of the *noumenal*. In Kantian terms, the noumenal dimension represents the boundary of what can be conceptually grasped; it is the "thing-in-itself" that underlies empirical phenomena but remains inaccessible to human cognition (B306-B310). Kant's transcendental philosophy confines itself to examining the conditions that make experience and knowledge possible without making claims about the metaphysical reality of universals or other entities independent of those conditions. By rooting his account in *a posteriori* conventions and empirical arrangements, Lewis replaces the transcendental focus on *a priori* conditions of possibility with a metaphysical framework that prioritizes empirical and systematic regularities.

From a Kantian perspective, one might argue that any theory addressing meaning or interpretation inherently engages with transcendental conditions, whether acknowledged or not. Kant coined the term "transcendental" precisely to emphasize the distinction between two different ways of establishing the foundations of judgments that exceed empirical limits. Transcendental logic grounds these foundations according to principles for judgment rooted in epistemic principles (Kant uses his theory of schemas), ensuring a critical awareness of the conditions of cognition. In contrast, dialectic—or what Kant calls the "logic of illusion"—grounds such principles on deceptive inferences, antinomies, paralogisms, and the reification of ideals, such as that of a "Supreme Being". While the finer details of this distinction go beyond our immediate scope, Kant's classification provides a valuable framework for understanding how the semantics of natural language, as proposed by Davidson and Lewis, can be situated within a transcendental problem space. Kant underscores the difference between a reflective, theoretical awareness of the conditions of signification and a dogmatic, uncritical attitude that takes these conditions for granted:

> Nevertheless there is something so seductive in the possession of an apparent art for giving all of our

> cognitions the form of understanding, even though with regard to their content one may yet be very empty and poor, that this general logic, which is merely a canon for judging, has been used as if it were an organon for the actual production of at least the semblance of objective assertions, and thus in fact it has thereby been misused. Now general logic, as a putative organon, is called dialectic (A61/B85)

This distinction is essential for identifying the potential risks in Davidson's and Lewis's frameworks.

Kant refers to dialectic as the misuse of transcendental reflection, wherein it is employed to ground logical theories that amount to little more than an *organon* of empty formulas—formal structures that lack substantive content. These formulas become the basis for illusions of pure reason when they are mistakenly treated as things-in-themselves or as judgments about the suprasensible realm. Such misuse occurs when theoretical constructs are taken to possess meaningful content, leading to speculative claims about entities or realities beyond what can be experienced or known through the senses.

When applied to the semantic theories of Davidson and Lewis, this helps illuminate that their respective frameworks arguably depend on foundational principles that, if uncritically assumed, could align with the kind of dialectical misstep Kant cautions against. Both theorists claim methodological neutrality, yet their reliance on implicit assumptions about meaning, judgment, and coherence places them within a transcendental problem space that they do not fully address. In the case of Lewis, his ontology of possible worlds exemplifies a move that explicitly bypasses epistemic conditions of judgment. By treating possible worlds as real entities that supervene on empirical arrangements, Lewis appears to construct a metaphysical framework that leans toward a realm of supraempirical entities. This naturalized metaphysical commitment runs counter to Kant's insistence that cognition be grounded in *a priori* epistemic conditions that frame our capacity to judge. As Kant argues in the *Critique of Pure Reason*, speculative constructs like the "world in itself" exceed the legitimate boundaries of human cognition and risk becoming illusions of pure reason when taken as objects of knowledge rather than heuristic tools (A644/B672-A645/B673).

Davidson, on the other hand, appears to avoid metaphysical commitments by grounding his theory in Tarski's truth-conditional semantics and a principle of charity that assumes shared rationality among speakers. However, this reliance on shared norms and observable behavior as sufficient for meaning reveals an implicit dependence on epistemic preconditions that Davidson does not fully explicate. As Kant notes, an uncritical reliance on such preconditions risks collapsing reflective inquiry into a dogmatic framework that takes these conditions as given, rather than critically examining their origin and limits. Davidson's approach, while ostensibly neutral, thus mirrors the kind of dialectical error Kant critiques: it presupposes coherence and rationality as universal without accounting for the transcendental conditions that make these assumptions possible.

It is in this latter sense that both authors cannot avoid being situated within a transcendental framework. In this sense, *not* reflecting on the underlying conditions of meaning results in a dogmatic approach that takes them for granted, failing to critically examine the assumptions that ground their theories. Thus, even though Davidson and Lewis reject transcendentalism in its explicit form—namely, as epistemic standards that define the boundaries of meaning—they nonetheless fail to maintain true transcendental neutrality.

7. The illusion of neutrality: transcendental dependencies in semantic frameworks

Despite our repeated assertion that these authors cannot escape a certain transcendentalism, this claim might appear innocuous and a superfluous insistence unless we demonstrate how this supposed commitment to neutrality ultimately spills over into other problems—problems that, in one way or another, belong to the sphere of the transcendental problematic. This is why our critique does not merely aim to highlight the presence of transcendental assumptions but to show how these assumptions generate significant philosophical consequences within the frameworks of Davidson and Lewis. Our critique does not deny that Davidson and Lewis can maintain the project of neutrality within their theories. It is entirely possible to sustain a truth-based semantic framework that seems indifferent to normative or epistemic presuppositions. However, the issue lies in the consequences of adhering to this illusion of neutrality. By neglecting to critically engage with the

underlying normative and epistemic conditions of their frameworks, Davidson and Lewis incur two significant costs.

A theory of truth that reduces meaning to a contentless agreement—where agreement is defined solely as the alignment of truth values within shared extensions—falls short of providing a robust account of meaning. This limitation points us toward the transcendental problem at the heart of such frameworks. When formal semantic theories encounter indecision, problematicity, or ambiguity, they cannot independently resolve these issues. Instead, they default to paradigmatic epistemic standards—normative structures derived from the dominant categories, judgment practices and methodological practices of a particular historical and cultural context. These standards, which determine what is rational, coherent, or meaningful, operate as the unacknowledged scaffolding for resolving interpretative disputes.

The dependence on epistemic paradigms means that formal semantic theories are not self-sufficient. They are always absorbed into or reformulated according to the prevailing epistemic framework. The failure of formal mechanisms based on semantic exhaustion to fully account for meaning without invoking epistemic standards leads us to a transcendental problematic. This arises because resolving semantic indeterminacy—by depending on the preconditions for meaning-making—can approach these preconditions either consciously or uncritically and dogmatically. In both cases, however, the issue remains fundamentally transcendental: either because there is an awareness of the conditions of unity (Kantian apperception; *KrV* B138) that underpin epistemic practices of judgment, or because the lack of such awareness results in a tacit and unreflective acceptance of the dominant standard for establishing this epistemic unity, thereby becoming captive to the categories and conceptual schemes imposed by the prevailing paradigms of the time.

Because of that self-insufficiency, these theories inadvertently reveal a limitation that invites the inclusion of a transcendental reflection. While these approaches aim for precision and universality, they reduce meaning to compatibility with formal systems or truth values, neglecting the interplay between meaning and the evolving standards of cognitive access and epistemic justification. As discussed in the introduction, the contrast between Lamarckian inheritance and Darwinian evolution illustrates this dynamic. Lamarck's view, where acquired traits are heritable, assumed direct environmental influence, while Darwinian

evolution, refined by Mendelian genetics, emphasized natural selection and random variation. These shifts not only redefined biological knowledge but also transformed the epistemic standards for evaluating it. In ignoring those epistemic standards, Davidson and Lewis fall short of accounting for the ways in which meaning is shaped by the shifting paradigms that govern what we consider valid knowledge or belief at any given time.

By reducing meaning to compatibility with formal systems or truth values, these theories operate within narrowly defined boundaries, ignoring the underlying conditions that make such systems possible in the first place. These transcendental conditions—such as the pre-structured forms of cognition, the intersubjective grounds of communication, and the norms governing the validity of knowledge—are what allow meaning-making to occur at all. The transcendental approach, rooted in traditions like the philosophy of Kant or Husserl (1931), addresses precisely these foundational questions: what are the conditions that must be in place for meaning to emerge?

This reliance on hidden preconditions manifests as a malady—a lack of awareness rooted in the absence of a remedy that addresses the deeper foundations of our meaning-making and communicative practices. We treat the symptoms while the underlying cause remains obscured, failing to ask the right questions about the foundation on which our language and meaning are built. First, this manifests through the normative standards that underpin shared rationality and coherence, as seen in Davidson's principle of charity and triangulation. Second, through the implicit rationalization processes in Lewis's account of linguistic conventions, where meaning is stabilized by communal practices that presuppose trust, truthfulness, and epistemic reliability. In both cases, the absence of direct awareness regarding the substance of the divergences in meaning or the content of each web of beliefs does not negate these divergences or the content of semantic agreements. Instead, it perpetuates them, often unconsciously, in alignment with historically entrenched epistemic paradigms, which thereby substitute for unexamined consensus.

Conclusion

Davidson and Lewis—two of the most ambitious contributors to natural language semantics—ironically find themselves entangled in challenges at the very boundaries their frameworks were designed

to transcend. Both rely heavily on truth-conditional semantics and systematic compositionality, aiming to provide a neutral and universal foundation for interpreting meaning. However, these approaches are revealed to possess significant vulnerabilities upon closer examination.

Let us, then, survey the aftermath. Davidson, a bold navigator of linguistic terrains, endeavored to anchor meaning to the Tarskian framework of truth. Epistemic agreements are therefore treated as secondary or even irrelevant, reducing the richness of meaning to its manifestation in observable linguistic patterns. His leap from truth to meaning represents an ambitious attempt to provide a scientific and systematic solution to the long-standing complexities of semantics. Yet, this reductionist approach trades the depth of meaning—shaped by cultural and epistemic paradigms—for a streamlined and operationally effective model. This principle, assertive in its simplicity, compresses the expansive complexity of human meaning into observable patterns of assent and dissent, leaving the deeper epistemic dimensions unexplored.

Lewis, equally ambitious, envisioned a theory of meaning rooted in the vast landscape of possible worlds. His reliance on conventions, underpinned by trust and truthfulness, offered the promise of a linguistic order free from ambiguity. Yet his framework, despite its appeal to universals and systematic rigor, ultimately falters. Lewis, too, is drawn to the lure of reductionism, reducing meaning to a matrix of truth values while neglecting the epistemic paradigms that underpin any arbitration between rival meanings when indeterminacy arises. This aligns both Lewis and Davidson with Gross's principle of exhaustion, which ensures that their frameworks are descriptively and normatively minimalist, focusing solely on truth-conditional structures as sufficient for capturing meaning. There is nothing beyond this to distinguish between them, as meaning is reduced to the formal congruence of truth values. In semantic terms, one could argue that they are essentially equivalent in terms of what is considered meaningful.

We then observed that the cost of this semantic exhaustion is a contentless agreement and the self-insufficiency of semantics. By reducing meaning to *contentless agreement*—a formal congruence of truth values—Davidson and Lewis construct theoretical frameworks that amount to skeletal approximations to the rich, dynamic interplay of meaning and context. This reductionist approach overlooks the epistemic paradigms and the intricate conditions that shape the arbitration of meaning. Such paradigms are not static backdrops but active frameworks that guide

interpretation, mediate between rival meanings, and anchor language in the shifting sands of historical, cultural, and cognitive contingencies. By declaring the epistemic divergences coming from different epistemic paradigms or different webs-of-beliefs (or conceptual schemes) irrelevant or incoherent, the theory flattens the complexity of meaning into a shallow veneer of agreement—a neat but hollow margin of truth, where compatibility of truth values alone reigns supreme. Davidson and Lewis, despite their contrasting approaches, seem united in this grand reduction: meaning, for all its supposed nuance, is crushed into tidy patterns of agreement, limited by rules as learnable as they are devoid of depth. Without acknowledging these deeper structures, their theories risk abstracting meaning away from the very foundations that make it intelligible and actionable.

This observation resonates deeply with the project of incorporating natural language into the Tarskian framework, where the limitations of formal methods become evident. Tarski's own skepticism toward the idea of capturing meaning solely through formal systems should have been more fully respected. In natural language, the conceptual foundations of scientific paradigms, the historical consensuses that shape shared understanding, and the normative institutions underpinning patterns of practice expose disagreements, conflicts, and incommensurabilities that cannot be neatly resolved by formal methods alone. The disagreements, conflicts, and incommensurabilities reflect the content or substance of a worldview, or a form of agreement, that is far from empty. These issues often manifest as conflicts between scientific paradigms, where the lack of shared understanding deepens the challenges in reconciling differing perspectives. Davidson's and Lewis's theories, deliberately committed to discarding attention to this dimension, ultimately risk flattening the rich complexity of meaning into formal structures that ignore the depth of interpretive and epistemic variability.

The final step of our argument demonstrates that Davidson and Lewis are situated within the realm of a transcendental debate. For Kant, this realm is not limited to those who explicitly adopt a theory of categories or overtly commit to epistemic principles and principles of judgment. It also encompasses those who assume a dialectical position—one that implicitly engages with transcendental conditions but does so uncritically, resulting in what Kant identifies as a logic of illusion.

This logic of illusion arises when foundational assumptions are treated as neutral or self-evident, bypassing critical reflection on the

conditions that make them possible. In the case of Davidson and Lewis, their theories' reliance on semantic exhaustion reveals this tension. Semantic exhaustion reduces meaning to a contentless agreement. Such a framework, while ostensibly neutral, depends on unacknowledged epistemic and normative paradigms for resolving ambiguities and disputes. This dependency situates their work within a transcendental problem space, even if they explicitly deny such commitments.

It is important to note that the apparent neutrality of Davidson's and Lewis's frameworks, and their deliberate avoidance of transcendental reflection, does not mean that the vacuum created by this indifference remains unoccupied. On the contrary, their theories implicitly rely on substantive normative ideals, which compromise their claimed indifference to the transcendental conditions of meaning. The neglect of these transcendental assumptions has already been shown to manifest in two key ways: (1) the reduction of meaning to a formal structure that fails to bridge the gap between contentless agreement and a comprehensive theory of meaning, and (2) the uncritical limitation of meaning, shaped by the dominant methodologies and judgments of a specific historical period.

This becomes especially clear when we examine key elements of Davidson's and Lewis' theories, which, despite claims to neutrality, implicitly depend on substantive normative ideals that are not neutral. Davidson's principle of charity mandates interpreting speakers as largely coherent, truthful, and rational. This normative ideal is not an empirical observation but a regulative assumption about what counts as meaningful communication. By grounding semantic interpretation in the epistemic expectation of belief coherence, Davidson implicitly relies on a transcendental stance that posits conditions of possibility for meaning. These conditions—rationality, coherence, and truth-seeking behavior—are not neutral. Instead, they are historically and culturally mediated, shaped by prevailing norms of epistemic adequacy. Lewis's theory of conventions similarly depends on shared norms of trust and truthfulness, which are essential for the stability of linguistic practices. However, these norms are not purely formal constructs; they are embedded within specific sociohistorical contexts. The very idea of convention presupposes a shared background of cultural and epistemic practices, which define what counts as "trustworthy" or "truthful."

Defending these concepts as mere *ceteris paribus* clauses—idealized limit cases that hold only under perfect conditions—is insufficient. In

practice, these assumptions align with dominant linguistic practices or prevailing "language games," which ultimately define the standards for agreement and coherence. This alignment reveals that both Davidson's and Lewis's frameworks are shaped by unacknowledged epistemic conditions, undermining their purported neutrality. Far from escaping the transcendental problem space, their theories inhabit it implicitly, demonstrating that the indifference to transcendental reflection does not eliminate its influence—it merely obscures it.

This further reinforces the idea that a theory of truth, which reduces meaning to a contentless agreement—defined solely by shared truth margins in extension, where the formal alignment of truth values is sufficient to determine agreement without addressing the richer, substantive aspects of meaning—fails to serve as a robust theory of meaning. Meaning, in this sense, is an event in natural language and in other contexts of our practical life. In our lives, what we perceive as consistent, regular, and predictable in language patterns emerges from sources of normative imposition, assumptions about how to live, how to regulate society, the language games we engage in, and other such frameworks. The formalization of these patterns follows later. By prioritizing formalization, natural language semanticists reverse the proper order of priorities, bypassing the opportunity to examine these foundational conditions from a reflective perspective.

References

Carnap, R. (1947). *Meaning and Necessity: A Study in Semantics and Modal Logic*. University of Chicago Press.

Carnap, R. (1963). *Formalization of Logic*. Harvard University Press.

Davidson, D. (1973). On the Very Idea of a Conceptual Scheme. In D. Davidson, *Inquiries into Truth and Interpretation* (pp. 183-198). Oxford University Press. https://doi.org/10.2307/3129898

Davidson, D. (2001). *Inquiries into truth and interpretation*. Clarendon Press. https://doi.org/10.1093/0199246297.001.0001

Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, 100, 25-50.

Gross, S. (2012). Davidson, First-Person Authority, and the Evidence for Semantics. In P. Kotatko, H. P. Martin, & P. Pagin (eds.), *Interpreting Davidson* (pp. 233-252). Stanford University Press. https://doi.org/10.1093/acprof:oso/9780199697519.003.0011

Husserl, E. (1931). *Ideas: General Introduction to Pure Phenomenology.* W. R. Boyce Gibson (trans.). George Allen & Unwin.

Kant, I. (1998). *Critique of Pure Reason.* P. Guyer & A. W. Wood (trans.). Cambridge University Press.

Kaplan, D. (1964). *Foundations of Intensional Logic* [Doctoral dissertation, University of California, Los Angeles]. UCLA Library.

Kripke, S. (1963). Semantical Considerations on Modal Logic. *Acta Philosophica Fennica*, 16, 83-94.

Lewis, D. (1969). *Convention: A Philosophical Study*. Harvard University Press.

Lewis, D. (1970). General Semantics. *Synthese*, 22(1/2), 18-67. https://doi.org/10.1007/BF00413598

Lewis, D. (1975). Languages and Language. In K. Gunderson (ed.), *Minnesota Studies in the Philosophy of Science. 7. Language, Mind, and Knowledge* (pp. 3-35). University of Minnesota Press.

Lewis, D. (1999). New Work for a Theory of Universals. In J. Kim & E. Sosa (eds.), *Metaphysics: An Anthology* (pp. 283-300). Blackwell Publishing.

Lepore, E. & Ludwig, K. (2012). Introduction: Davidson's Philosophical Project. In G. Preyer (ed.), *Donald Davidson on Truth, Meaning, and the Mental* (pp. 1-36). https://doi.org/10.1093/acprof:oso/9780199697519.003.0001

Lepore, E. & Ludwig, K. (2013). Truth in the Theory of Meaning. In K. Ludwig & E. Lepore (eds.), *A Companion to Donald Davidson* (pp. 175-190). Wiley-Blackwell. https://doi.org/10.1002/9781118328408.ch10

Montague, R. (1960). Logical Necessity, Physical Necessity, Ethics, and Quantifiers. *Inquiry*, *4*(1-4), 259-269. https://doi.org/10.1080/00201746008601312

Montague, R. (1968). Pragmatics and Intensional Logic. *Synthese*, *22*(1/2), 68-94. https://doi.org/10.1007/BF00413599

Montague, R. (1971). The Proper Treatment of Quantification in Ordinary English. In J. Hintikka, J. Moravcsik, & P. Suppes (eds.), *Approaches to Natural Language* (pp. 220-247). Reidel.

Quine, W. V. O. (1969). *Ontological Relativity and Other Essays*. Columbia University Press. https://doi.org/10.7312/quin92204

Ramberg, B. T. (1989). *Donald Davidson's Philosophy of Language: An Introduction*. Blackwell.

Rattan, G. S. (2005). Davidson, Semantic Deflationism, and Dummett's Dilemma. *Iyyun: The Jerusalem Philosophical Quarterly*, 54, 39–63.

Rorty, R. (1999). Davidson between Wittgenstein and Tarski. *Crítica: Revista Hispanoamericana de Filosofía*, *30*(88), 49-71. https://doi.org/10.22201/iifs.18704905e.1998.696

Scott, D. (1970). Advice on Modal Logic. In K. Lambert (ed.), *Philosophical Problems in Logic* (pp. 143-173). Reidel. https://doi.org/10.1007/978-94-010-3272-8_7

Tarski, A. (1944). The Semantic Conception of Truth and the Foundations of Semantics. *Philosophy and Phenomenological Research*, *4*(3), 341-376. https://doi.org/10.2307/2102968

Tarski, A. (1956). The Concept of Truth in Formalized languages. In A. Tarski, *Logic, Semantics, Metamathematics* (pp. 152-278). Clarendon Press.

Wittgenstein, L. (2009). *Philosophical Investigations.* P. M. S. Hacker & J. Schulte (trans.). Wiley-Blackwell.