ADVANCE SOCIAL SCIENCE ARCHIVE JOURNAL
Available Online: https://assajournal.com
Vol. 04 No. 01. July-September 2025.Page#.2202-2217
Print ISSN: 3006-2497 Online ISSN: 3006-2500
https://doi.org/10.55966/assaj.2025.4.1.0117
Platform & Workflow by: Open Journal Systems

Bridging Philosophical Foundations and Computational Realities: Semantic Under-Specification from Frege to Large Language Models

Tahir Qayyum
University of Azad Jammu & Kashmir, Muzaffarabad
tq.khan@hotmail.com
Anam Tahir
Higher Education Department, AJ&K
anum_rafique@hotmail.com
Iftikhar Ahmed Shaheen
Majmaah University, KSA
I.shafi@mu.edu.sa
Syed Atif Amir Gardazi
Majmaah University, KSA
s.gardazi@mu.edu.pk

ABSTRACT

Semantic underspecification occurs when linguistic expressions carry partial meaning, requiring context for full understanding. It poses key challenges across philosophy, cognitive science, and NLP. This review identifies five developmental stages: (1) Classical theories by Frege, Russell, and Davidson created truth-conditional frameworks but encountered difficulties with indexical and belief contexts due to assumptions of full specification; (2) Formal models like QLF, MRS, and Hole Semantics introduced computational underspecification to handle structural ambiguities such as quantifier scope; (3) Cognitive studies show humans use underspecification strategically for efficiency, relying on pragmatic inference and semantic memory; (4) Hybrid neuro-symbolic models like UMR and Glue Semantics combined structural ambiguity resolution with neural inference but lacked uncertainty modeling; (5) Modern NLP research highlights gaps: LLMs can detect underspecification but often overcommit to deterministic interpretations, and multimodal systems do not effectively utilize context. Cross-linguistic entropy models suggest grammatical underspecification as a strategy for cognitive efficiency. To bridge human semantic flexibility enabled by incremental processing and pragmatic co-construction—with computational systems, we propose integrated neuro-symbolic architectures incorporating explicit uncertainty modeling, multimodal grounding, and entropy-aware design. This approach paves the way for AI to achieve human-like language understanding.

Keywords: *Philosophical Foundations, Computational Realities, Semantic, Frege, Large Language Models*.

INTRODUCTION

Underspecifcation, in general, refers to linguistic ambiguities deliberately and purposefully made through concise representations to draw a parallel between potential interpretations (Lai

& Piñango, 2017; Manshadi et al., 2018). Disambiguating underspecified choices help prevent the overwhelming complexity that can come with trying to distinguish every possible meaning all at once (Szarvas et al., 2012; Chaves, 2003) by maintaining common semantic elements while trusting context or practical processes to resolve details (Pezzelle, 2023; Frisson, 2009). The cognitive significance of underspecifiatons is rooted in reflecting how human language prioritizes efficiency such as casual conversation often shares incomplete information, which helps streamline storage and processing (Pezzelle, 2023).

However, underspecifications and ambiguities, although interchangeable at times, are distinguished concepts. An ambiguity refers to discrete meanings such as at lexical level, (e.g., "bank" = [FINANCIAL INSTITUTION] or [RIVER EDGE]),

Underspecification, on the other hand, provides a unified and context dependent schematic meaning (Pezzelle, 2023; Lesmo et al., 2007). An example such as "book" activates a general sense ([INFORMATIONAL OBJECT]) initially, with specific senses ([TEXT] / [RESERVATION]) accessed only when contextually acceptable (Lesmo et al., 2007).

Underspecification manifests in various forms, including default grammatical encodings such as the pronoun "they" and gender inflection, structural ambiguity (Renkema, 2004), omitted scopal relations, or inference dependent on context. Unlike ambiguity, underspecification does not necessarily produce multiple interpretations; instead, it deliberately leaves semantic gaps that require contextual, inferential, or perceptual cues to be adequately filled (Egg, 2019).

This review delineates the progression of underspecification research, beginning with classical semantic theories (Frege, Russell, Davidson), advancing through computational frameworks (Hole Semantics, MRS, Dominance Constraints), and culminating in cognitive and psycholinguistic investigations. It concludes with contemporary multimodal NLP and large model research (e.g., Pezzelle 2023; Wildenburg et al., 2024).

HISTORICAL FOUNDATIONS (PRE-2000)

Fregean Semantics: Sense, Reference, and Cognitive Significance

Frege's (1982) revolutionary distinction between sense and reference substantially transformed our conceptual understanding of meaning by delineating Sinn (sense) from Bedeutung (reference). Frege's seminal distinction between sense (Sinn) and reference (Bedeutung) is a cornerstone of philosophical semantics (Oliveira, 2015; Textor, 2019; Soutif, 2022). Frege argues that expressions, particularly proper names and complete assertive sentences, possess both a sense and a reference (Oliveira, 2015; Farzam, 2023). This refers to the object or entity that an expression points to in the world (Oliveira, 2015; Kotzé, 2001). For instance, the reference of the name "Aristotle" is the historical philosopher himself. For sentences, Frege controversially proposed that their reference is their truth-value (the True or the False) (Eder, 2019; Martin, 2006). This idea was driven by his Principle of Composition, which suggests that the reference of a complex expression is determined by the references of its constituent parts (Eder, 2019). While, Sense is the "mode of presentation" or the way in which the reference is given (Farzam, 2023; Kotzé, 2001). Two expressions can have the same reference but different senses. A classic example is "the Morning Star" and "the Evening Star." Both refer to the same celestial body (Venus), but they present it in different ways, leading to distinct cognitive significance (Penner, 1995). Frege posited that sense captures the cognitive value and is distinct from mere subjective ideas or psychological associations (Soames, 2010;

Kotzé, 2001). For sentences, the sense is the thought or proposition expressed, which determines its truth-value (Martin, 2006).

Frege's system assumes highly detailed semantics, where each term has a full sense and reference. However, this idealised view clashes with everyday language, which often leaves out explicit scope, pronoun antecedents, or ellipses. In these cases, Frege's framework struggles to manage semantic under-specification because it expects all meaningful parts to be explicitly stated, even when some are deliberately omitted or postponed (Lai, & Piñango, (2017).

Russellian Descriptions: Structured Propositions and Referential Failure

Russell, originally influenced by Frege, diverges considerably later, especially in his analysis of definite descriptions and theories of meaning (Oliveira, 2015; Neale, 2001). His famous theory of descriptions, outlines in "On Denoting," explains how phrases like "the current King of France" can remain meaningful even if they lack an actual referent. Russell argued that definite descriptions are not proper names referring to objects but are "incomplete symbols" that are analyzed away in a logical paraphrase (KRIPKE, 2008).

For Russell, the meaning of a proper name is its referent (direct reference theory). In contrast, definite descriptions derive their meaning from their role in establishing the truth-conditions of the sentences in which they appear, rather than by directly referring to an object (Penner, 1995). This methodology sought to address issues such as sentences involving non-existent entities and the problem of informative identities, without resorting to Fregean senses. Russell's work highlighted the importance of the logical structure of sentences and their connection to facts in the world (Soames, 2010).

Russell's rejection of a Fregean sense for descriptions suggests that a sentence's meaning does not rely on its parts having a "sense" as Frege believes, but on their capacity to denote or their contribution to the proposition's truth value (Oliveira, 2015). His naturalistic epistemology also supports a view of knowledge based on experience and logical reasoning, impacting later empirical studies of language (Hedger, 2013; Vasile, 2013).

His structured proposition model enhances logical clarity by translating noun phrases into quantifier expressions. Nonetheless, Russellian semantics encounter difficulties in attitude contexts; for instance, sentences like "Smith believes that cordates are renates" involve empty names or unfamiliar terms but still convey meaningful content. Russell's methodology would regard these as false or nonsensical, neglecting the potential for a speaker to genuinely believe something despite a lack of knowledge about the referent. Contrary to Frege's notion of cognitive sense, Russell's formalism does not effectively address belief opacity or the differentiation between de re and de dicto attitudes.

Davidsonian Semantics: T-Sentences and Truth Conditions

Davidson supports a semantic program based on Alfred Tarski's truth theory (Sabir, 2023; Soames, 2014; E.G. & Opande, 2021). He argued that a natural language's meaning could be explained through a corresponding truth theory (Sabir, 2023; Stanley, 1996). Davidson extends semantic theory by introducing T-sentences, such as' "Snow is white' is true if and only if snow is white,' to formalise a truth-conditional perspective of meaning. Although this methodology effectively captures compositional semantics, it offers limited guidance in interpreting phenomena such as polysemy, ellipsis, pronouns, or modal variation. The model tends to condense multiple interpretations into a single truth value, thereby neglecting distinctions that impact pragmatic understanding and inferential nuances.

Earlier propositional theorists focused exclusively on declarative sentences to assign truth values, whereas more contemporary theorists include questions, commands, and culturally unique utterances (Griesel, 2007; Su et al., 2020). Modern discussions about meaning seem always to have some underlying philosophical currents that extend beyond truth conditions and references (Jackson, 2020; McElvenny, 2018). These issues pose a challenge to classical semantic theory, demonstrating outdated assumptions in contemporary theory.

EMERGENCE OF UNDERSPECIFICATION FORMALISMS: A NEW PARADIGM

Languages naturally include ambiguities like scope, anaphora, and ellipsis. Traditional semantic theories tackle these problems in deep detail for all interpretations, which is, for sentence complexity, an enormous problem for computational tools (Egg et al., 2001). With the help of computational tools, these minor ambiguities can lead to thousands of different forms. Alshawi and Van Eijck (1991), dealing with this problem, propose Quasi-Logical Form (QLF) and Underspecified Logical Form (ULF) where scope-ambiguity is preserved by halting the process of deciding the order of quantifiers, in other words, acting as intermediate forms between sentence and full disambiguation. These two frameworks include quantifiers and/or pronouns without assigning their referents and, or scope too early.

In order to model the hierarchy of relationships effectively, Bos (1995, 1997) introduced Hole Semantics which uses labelled "holes" along with dominance constraints. Instead of having to compute all scope combinations in advance, this method progressively resolves the scope by filling the holes at a later stage of computation. Niehren and Koller (2001) and Arts et al. (2011) show that dominance constraints form the basis of frameworks like MRS (Minimal Recursion Semantics). Copestake et al. (1995) also created MRS to decompose semantics into elementary predications (EPs) with scope-controlled semantics, creating dominant graphs efficiently. Constraint Language for Lambda Structures (CLLS) Koller et al. (2000) combines dominance and parallelism in explaining how ellipsis, anaphora and scope relate to one another.

The method again enhanced its use for complicated phenomena (Sag & Wasow, 2014; Nicoletti et al., 2023). This set of tools provides rigorous methodology for eliminating syntactic ambiguity along with facilitating the theoretical semantic modals (Sanford & Strut, 2002; Fecher & Steffen, 2005).

Not all semantic phenomena fit existing frameworks, such as referential underspecification, ellipsis, implicit content, and pragmatic inference, which go beyond structural ambiguity. Formalisms like MRS or Hole Semantics cannot address these internally and need external mechanisms. These representations are formal and static, unable to model how readers use context and inference to resolve ambiguity incrementally, making them disconnected from pragmatic viability. Additionally, technical complexity exists. Systems like CHORUS, called the “Swiss Army Knife of Underspecification,” help convert formalisms and solve dominance graphs, but full semantic resolution still requires advanced symbolic tools and grammar constraints.

However, Koller (2004) and Ebert (2005) observe that these expressive frameworks mainly address structural underspecification, such as quantifier scope, but fall short in managing Referential vagueness, Ellipsis, and Implicit inference. They also lack in handling coreference, pragmatic defaults, and culturally grounded inference, which are vital for natural language understanding. Researchers have tried to improve underspecification frameworks by using dominance constraints or labelled parallelism. However, even these strict rules cannot cover all forms. For example, NDCS can show that a quantifier should dominate a negation within particular scopes. However, it cannot specify that an omitted pronoun may have a human or non-human antecedent depending on context.

This structural reinforcement improves tractability; however, it does not resolve the issue that various forms of ambiguity, such as belief contexts, deixis, and pragmatic inference, cannot be simplified to structural tree ambiguity. These types of meanings require context-sensitive resolution, which purely symbolic systems are incapable of managing.

Fregean, Russellian, and Davidsonian semantics serve as foundational frameworks for meaning theory. However, they assume fully detailed semantic content, which conflicts with the inherently ambiguous and context-dependent nature of natural language. Formal underspecification frameworks represent a major methodological improvement by enabling meanings to stay deliberately underspecified, postponing their resolution until the context or computational inference clarifies them. This approach balances compositionality with interpretive flexibility.

Nevertheless, symbolic models are solely insufficient in capturing contextual resolution, pragmatic inference, or human processing dynamics. They concentrate on structural underspecification, such as quantifier scope, yet do not confront referential or pragmatic underspecification. Significantly, they establish the groundwork for subsequent cognitive, psycholinguistic, and neuro-symbolic approaches, where interpretive nuance is crucial for comprehending meaning within context.

PSYCHOLINGUISTIC AND COGNITIVE PERSPECTIVES (2000–2015)

This approach contrasts with traditional views that assume complete analysis of all linguistic levels during comprehension (Sanford & Graesser, 2006). From a cognitive-linguistic perspective, researchers have long understood that speakers tend to simplify meaning construction, often employing semantic underspecification as an efficiency strategy. Ambiya (2008) and Lin et al. (2014) propose that speakers intentionally omit referents, pronouns, or deixis, particularly with regard to gender and number, because the surrounding context typically supplies sufficient information to fill in the gaps. This concept aligns with Tulving's (1972) semantic memory model, which differentiates between episodic and semantic memory: the latter being a adaptable, context-independent repository of general knowledge that allows language users to access meanings even when surface forms are incomplete or ambiguous.

Tulving (1972) expands on the encoding specificity principle by highlighting that memory retrieval depends on the overlap between encoding and retrieval contexts. Semantically, speakers encode partial meanings within specific contexts and fully access or infer them during retrieval. This accounts for why native speakers can easily understand underspecified pronouns or ellipsis: their mental lexicon retrieves or infers meanings from previous knowledge without taxing working memory.

Researchers (Frisson, 2009; Frisson & Pickering, 2001; Devitt, 2021) demonstrate that polysemous words (such as "bank", "court") activate a wide range of meanings, making it complicated to indicate one directly. However, when filtered through the initial step, context determines the correct interpretation. MacGregor et al. (2020) also reveal, through MEG/EEG studies, that resolving these semantic ambiguities involves several brain processes. It begins with initial activation, followed by selecting the suitable meaning, and eventually reinterpreting the sentence if necessary.

While investigating challenges faced by L2 learners through eye-tracking, Zhao (2024) asserts that English speakers learning Spanish exhibit notably reduced accuracy when dealing with non-standard word orders. He emphasises that structural complexity and underspecification impede accurate processing, leading to comprehension. Complementary areas in the brain are involved in language comprehension (Blackett et al., 2022). For instance, the anterior temporal lobe assists in disambiguation, whereas the posterior temporoparietal cortex manages thematic relationships. This interconnected neural network supports unified semantic processing. Rasin and Aravind (2020) contend that children develop sensitivity to semantic nuances early, including underspecification patterns, as demonstrated in studies of quantifier acquisition such as "every."

Frisson & Pickering (2001) comment on ongoing debates about cognitive processes in activating and clarifying underspecified meanings, highlighting significant research gaps. They emphasise the importance of examining how linguistic context, world knowledge, and individual differences influence this process. Additionally, they underline the role of prior experiences and semantic retuning, as exposure changes word interpretations and ultimately impacts how underspecification is processed (Gilbert et al., 2021).

While cognitive linguistics emphasizes universal concepts (Mirdehghan, 2008; Le, 2024), the manifestation and management of underspecification vary across languages, making universal models difficult. Additionally, the influence of culture, personal experience, and the systematic nature of semantic change require further study (AlBzour, 2016; Panther & Thornburg, 2023).

Although today's NLP models are impressive, they still struggle with subtle, context-dependent features of human language, especially with implicit or underspecified information (Wildenburg et al., 2024; Mao et al., 2023). Addressing this gap requires more sophisticated models that incorporate broad world knowledge and perform deeper semantic and pragmatic reasoning.

Pragmatic Inference and Speaker Intent

Language underspecification is often a deliberate choice by speakers to attain their communicative goals. Based on Gricean implicature and Rational Speech Act (RSA) theories (Frank & Goodman, 2012; Goodman & Frank, 2016), speakers tend to use ambiguous phrases which are tacitly modified based on context, for example, claiming "The artist finished it" even without overt reference to "it," as the hearer can infer the referent based on shared knowledge.

RSA models codify this reasoning: a literal listener weighs potential interpretations, decides which are plausible in the context, and the speaker expects this. Pragmatics thus completes the gaps that grammar or logic cannot fill by itself. It follows that underspecification arises from shared expectations, cultural practices, and the inferential ability of the listener. It also implies that meanings are constructed collaboratively, and the speaker leaves space for interpretation and the listener infers the intended meaning implicitly, characteristic of natural conversation.

Psycholinguistic Observations: Eye-Tracking and Neural Evidence

The most important point to note in ERP research is that readers very quickly resolve underspecified language using predictive inference, and that this could occur, and often does, before explicit contextual cues are introduced (Huth et al., 2016; Degen, 2019). For example, upon encountering ambiguous phrases (e.g. "bank" within the vicinity of a river), readers initially activate both meanings. However, then they use context to quickly (and often apparently pre-consciously) refine which meaning was more relevant. Additionally, Frisson

(2009) and Egg (2019) make an important distinction between ambiguity, or picking from available meanings (for example, "bat" about either an animal or a cricket tool), and underspecified meaning, or initially accepting incomplete meanings that become more established over time. Their findings suggest that readers seem to intentionally wait to interpret underspecified expressions until the context makes it important, which would also reduce the cognitive load from making a confirming 'commitment' early on in reading.

Cognitive Integration: Memory, Inference, and Efficiency

A broader cognitive framework is required to comprehend the nature of underspecification. A framework that may combine semantic memory and pragmatic cues to psycholinguistic inferences. Abstract schemas are stored in semantic memory, such as "cat", which usually does not entail a detailed inquiry to recognise the meanings. In such a scenario, meanings are independent of contextual representations and mark flexibility. Pragmatic inference, however, needs contextual representations, prior knowledge and expectations for the successful interpretation of partial meanings. The ERP and eye-tracking studies discussed in the previous section imply that readers, in real time, employ and integrate these processes for fluent understanding of partial processing. This significant psycholinguistic and cognitive model offers valuable insights to account for semantic underspecifications, although it, too, has certain limitations. For instance, they propose that semantic memory assists adaptability; however, empirical evidence suggests it is hardly influenced by context and episodic memory. As the systems associated with memory are interlinked, processing and interpreting partial meanings at times require episodic memory recall, particularly when contextual cues are limited or unfamiliar. Shared context and cooperative listeners are the characteristic features of pragmatic models, yet these assumptions may not be applicable in cross-cultural or otherwise situations and settings, which challenge the universality of underspecifications and a single and unified approach to resolve them. Psycholinguistic research often uses controlled syntactic structures, thereby making it complex to ascertain whether similar predictive inferences occur in natural discourse with complex ambiguities. These models often neglect the deliberately maintained vagueness as a strategy, particularly in politics,, where more integration of cognitive and pragmatic approaches is required.

In summary, from 2000 to 2015, there was a detailed psycholinguistic and cognitive exploration of semantic underspecification. Tulving's (1972) memory theories support the notion that meanings do not have to be fully defined during production. Pragmatic frameworks view underspecification as a useful communicative strategy rather than a flaw, and experimental findings show that readers quickly and effectively process and interpret partial or ambiguous expressions.

COMPUTATIONAL FRAMEWORKS AND SYMBOLIC–NEURO HYBRIDS (2015–2022)

The merging of computational frameworks and symbolic-neuro hybrid systems represents a key frontier in artificial intelligence. It aims to combine human-like reasoning, characterized by symbolic processes, with pattern recognition primarily driven by neural methods. According to Pezzelle (2023), although significant progress in deep learning has appropriately addressed the semantic processing, it still needs further innovations owing to the complex nature of natural languages. Advanced computational frameworks need to enhance the capabilities of existing large language models for comprehensive redressal of semantic underspecifications. Such frameworks need to blend symbolic and neural approaches (Pezzelle, 2023).

Mao et al. (2023) and Wildenburg et al. (2024) highlight semantic underspecification, such as expressions like "Don't spend too much," which lack certain details, viewing it as a clever way to make language more efficient. Rather than seeing this as a mistake, it actually helps our brains process information more easily by relying on context and clues.

Dalrymple et al.'s (1993) Glue Semantics offers a flexible approach to combining meanings using linear logic, making a clear distinction between meaning and strict syntax. Van Gysel et al. (2021, 2024) build on this idea with the Uniform Meaning Representation (UMR), which creates cross-linguistic meaning graphs that effectively preserve ambiguities, like unresolved quantifiers, by using abstract concept nodes.

Wildenburg et al. (2024) point out some key limitations with their DUST dataset: although LLMs can recognize underspecification when prompted, they often lean towards producing one fixed interpretation, like a single reading of "they." This approach differs from the flexible way humans think about language, and we see only a weak link between this behaviour and measures like perplexity. Pezzelle (2023) also reminds us that unimodal systems cannot quite match how people combine social and visual cues to make inferences.

Symbolic frameworks like Glue/UMR keep their structure intentionally ambiguous, but can be a bit limited practically. To create better solutions, we need probabilistic outputs that explore relevant domains, especially when inputs are underspecified. First, multimodal grounding helps us understand visual and contextual clues to clarify deixis. Second, cross-linguistic hybrid systems blend UMR graphs with neural inference modules, making language comprehension more robust. Lastly, benchmarks like DUST focus on minimal pairs to target and resolve specific ambiguities efficiently. This exciting shift toward neuro-symbolic integration truly highlights a significant change—focusing more on human-like flexibility rather than strict disambiguation. It is a step toward making technology more adaptable and intuitive for everyone.

EMPIRICAL STUDIES (2023–2025)

In an innovative position paper, Pezzelle (2023) reconceptualizes semantic underspecification as an asset - as opposed to a defect - in human communication. He provides pronouns such as "they" (which notably conceal the details of gender or number) as an excellent example, demonstrating each speaker's expectation that the listener can leverage multimodal clues - whether visual context clues, situational context clues, or context clues derived from shared cultural knowledge - to fill in the absence of explicit semantic specification. Pezzelle (2023) states, further, that all but bottomless, current NLP systems are often incapable of exhibiting this capacity to make inferences and understand context, to which unimodal transformers in particular are ill-suited because the modality is limited to text, not context, yielding typical misclassifications (e.g. gender), hallucinations, or bias.

Multimodal models are promising,, but Pezzelle (2023) notes that they largely continue to under-utilise visual context to assist in coreference resolutions, e.g. correctly associating "they" with a relevant entity in an image, and experience referential ambiguity when cue(s) are scant. This deficiency highlights the necessity for architectures that genuinely replicate human-like contextual integration. Pezzelle (2023) cautions that, in the absence of mechanisms to explicitly seek clarification or communicate uncertainty, multilingual and multimodal systems may face adverse outcomes, especially in circumstances involving identity sensitivity or safety-critical scenarios. His framework emphasizes the importance of enabling models to dynamically default, defer, or clarify their understanding, analogous to human interactive strategies.

Wildenburg et al. (2024) introduced DUST, a specialised dataset comprising minimally paired sentences that differ in their level of specificity, such as “Don’t spend too much” versus “Don’t spend too much cash. " Their research evaluates whether pretrained language models (LMs) are capable of (a) recognising underspecified inputs and (b) interpreting them without bias. Although recent models like LLaMA 2 frequently identify an utterance as underspecified when explicitly prompted, they tend to assign only a single, arbitrary interpretation. This results in a failure to maintain interpretive ambiguity when multiple readings are possible, thereby contradicting human semantic flexibility.

Significantly, Wildenburg et al. (2024) demonstrate that the uncertainty inherent in models remains unrealistically minimal, indicating that language models almost invariably commit to responses without demonstrating caution or acknowledging uncertainty. This observation contradicts theoretical expectations, which suggest that underspecification should result in probabilistic distributions rather than deterministic outputs. They argue that frameworks for evaluation incorporate uncertainty estimation as a primary measure of quality and evaluate models on the capability to suspend certain judgments instead of being prone to hasty conclusions.

Current cross-linguistic research suggests a formal framework connecting grammatical underspecification to entropy alignment between grammatical and semantic distributions of features. The fundamental idea is that if the entropy of semantic features, e.g., natural gender, is higher than that of grammatical features, e.g., grammatical gender categories, then the system will merge several semantic values into one grammatical category and so create underspecified forms.

Cross-linguistic corpus analyses reveal a pattern of systematic grammatical underspecification: inanimates tend to default to unmarked gender forms even when their meanings are quite disparate (e.g., Spanish La Mesa [feminine table] versus el libro [masculine book]). This is evidence of a memory-surprisal tradeoff, where languages only encode the most essential distinctions to save cognitive effort while improving communication. Entropy-based models beautifully capture this notion as selective feature compression. They demonstrate how grammatical structures purposely leave out redundant semantic information to maintain structure tractability, ensuring more efficient processing of language. Van Gysel et al. (2021, 2024) also illustrate by automatic UMR conversion of Czech and Latin corpora that underspecified representations remain invariant under various transformations of structures. Semi-automated pipelines deliver meaning graphs whose grain size varies with parser strength and annotation depth. Critically, the conversions do not eliminate, but maintain, underspecification (e.g., leaving scope-ambiguous quantifiers as abstract nodes). This confirms underspecification to be computationally beneficial, so there is a design principle: UMR systems should not eliminate multiple interpretations from nodes, avoiding premature disambiguation.

Empirical studies reveal key gaps, such as UMR capturing structural indeterminacy but lacking inference and uncertainty modeling; neural models identify underspecification when prompted but tend to impose a single interpretation, sacrificing human-like ambiguity tolerance (Wildenburg et al., 2024). Additionally, multimodal systems fail to replicate human referential grounding (Pezzelle, 2023), often struggling with context-sparse deixis.

Researchers in the future may develop hybrid neuro-symbolic frameworks combining symbolic ambiguity storage with neural contextual resolution. Evaluation benchmarks like DUST, with underspecified minimal pairs and uncertainty metrics, measure progress. Ensuring safety through uncertainty-aware outputs, such as entropy thresholds and clarification requests, is vital for high-stakes areas. Exploring cross-linguistic generalization with entropy-guided formalisms seeks to predict typological variation in underspecification, expanding understanding across languages.

Recent empirical research (2023–2025) highlights a clear gap between human interpretive skills and current NLP models regarding semantic underspecification. While language models can identify ambiguity when asked, they often default to a single, deterministic meaning and usually do not express their uncertainty. Multimodal models also lack methods for context-driven inference or clarification. Although formal grammar-based theories explain why language naturally encodes underspecification, they do not offer solutions. Future AI systems need to include uncertainty-aware inference, multimodal grounding, and symbolic underspecification to better mirror human semantic adaptability.

Research into semantic underspecification has advanced in four stages. Classical semantics, linked to Frege, Russell, and Davidson, assumed expressions had fully defined meanings but struggled with real-world phenomena like indexicals and scope ambiguity. In the 1990s, frameworks like QLF, MRS, NDCS, and Hole Semantics enabled computational representation of interpretive flexibility, especially regarding quantifier scope and vagueness. Later, cognitive and pragmatic theories showed humans often choose underspecified forms for efficiency, using context and inference, as RSA models demonstrate. Recent neural and multimodal system research (Pezzelle 2023; Wildenburg et al. 2024) shows that advanced language models and vision-language systems face challenges with underspecification, tend to ignore ambiguity, default to deterministic interpretations, and fail to use context or generate clarifications—despite prevalent underspecified information.

Symbolic frameworks provide a refined method for handling structural ambiguity, but they cannot address world-based inference, referential vagueness, or implicit meaning. In contrast, neural models tend to oversimplify semantics by translating meaning into point estimates, which overlooks interpretive uncertainty.

Wildenburg et al. (2024) note that language models (LMs) can recognize underspecified inputs but rarely show interpretive uncertainty. They tend to assert a single interpretation, even with multiple possibilities, contrasting with human cognition. Research links grammatical underspecification to cognitive limits, like default gender marking when semantic entropy exceeds grammatical distinctions. Yet, in computational models, such cross-linguistic understanding is underrepresented, particularly beyond English. Benchmarks such as Winograd and VQA do not have truly ambiguous cases. Models therefore excel in experiments but have trouble dealing with underspecification in real-world contexts.

Models must be tested with corpora like DUST, that contain underspecified minimal pairs by definition and need to be treated by uncertainty-aware evaluations, not just correctness. Language models should have interpretive hesitation features, like calibration-aware outputs or probabilistic distributions, instead of single-point answers. Symbolic approaches like UMR or Glue Semantics maintain structured underdetermination, whereas neural inference modules can provide context-anchored resolution. A blending of these methodologies will likely better replicate human interpretation.

Pezzelle (2023) points out that human language necessarily comprises several modalities; NLP models must include visual and contextual information to parse underspecified constituents such as pronouns or referents correctly. Entropy formalism and cross-linguistic information imply that typology-dependent patterns of grammatical underspecification exist. Models must be generalized to add typologically heterogeneous input data in a way that avoids English-centric blind spots.

Although symbolic approaches describe the theoretical workings of underspecification, they do not capture the dynamic, context-sensitive reasoning that humans exploit. Conversely, neural models can detect ambiguity but do not represent semantic indeterminacy. Neither approach completely captures the full human interpretive process.

The overconfidence of language models when faced with underspecified data highlights a disconnect between their training objectives maximizing likelihood—and effective real-world communication. These models should be developed to indicate uncertainty in cases of ambiguity. Recent research demonstrates that grammatical markers frequently condense semantic nuances because of cognitive constraints. However, most models are not built to replicate this process; instead, they assume that lexical forms inherently reflect full semantic resolution.

IMPLICATIONS AND FUTURE DIRECTIONS

Existing frameworks, such as Uniform Meaning Representation (UMR) and Abstract Meaning Representation (AMR), should adopt multi-path interpretative structures to support highly underspecified meaning graphs effectively. Recent UMR initiatives (van Gysel et al., 2024) demonstrate their ability for structured multilingual annotation while maintaining abstraction. Allowing nodes to have multiple valid interpretations, UMR can achieve a balance between structural conciseness and semantic flexibility, adhering to the concept of compositional ambiguity.

Recent research indicates that memory–surprisal trade-offs are responsible for systematic underspecification in grammar, such as gender marking defaults influenced by entropy limits. Incorporating these constraints into grammatical frameworks enables symbolic models to anticipate when and why underspecification occurs, particularly in languages with high semantic variability condensed into fewer grammatical categories.

Following the models like lossy context surprisal (Futrell & Levy, 2019), which address memory limitations in partial context, future research should investigate how and to what degree readers process sentences with intentional underspecification over time. Experiments measuring N400 amplitudes, such as surprisal scores, and eye-tracking can reveal how incremental memory loss affects predictive processing during interpretation. Additionally, empirical studies involving interviews and reading tasks explore how speakers and listeners clarify ambiguous expressions through requests, inference, or default meanings.

This matches with RSA frameworks (Frank & Goodman) and offers some not very deep insights about interaction strategies. Future benchmarks should be beyond standard datasets, where they contain minimal pairs like DUST, more where underspecified items are explicit. The evaluations should develop uncertainty-aware metrics. The models should show what they can recognise as ambiguity and what they can reason out of ambiguity, not what they can singularize as disambiguation. Per Pezzelle (2023), their models could incorporate vision and gaze tracking or context when confirming underspecified referents. It would be helpful if the neural architecture could use external sources of information and slowly be able to settle for clarification while understanding so that understanding across questions would flow and be accurate.

In general, language models are much more than simply giving a fixed response. By incorporating some methods, like threshold entropy, confidence intervals, or sampling different hypotheses across decoding and training time, they can demonstrate that they can better reflect when they are unsure/stuck or confused- capturing similarity to how humans understand and derive meaning (Wildenburg et al., 2024). Similar to this, when they receive incomplete or unclear inputs, models might likely cause misplaced biases or inaccuracies (like misgendered pronouns, hallucinations, misunderstandings and unclear inputs, etc.). When this happens, it would be useful for them to be able to recognize unclear contents, follow up, or come back to completely safe or to treat as a blessing.To handle this gracefully, systems should be able to spot vague content, ask for clarification, or fall back to safe options. Especially in important areas like healthcare, law, or social media moderation, it's essential to include disclaimers or ask for user consent to keep things transparent and trustworthy.

Theoretical models ought to evolve in order to include not just structuring meaning, but also to accept ambiguity as a significant part of communication. Cognitive science should reflect on how people read or employ incomplete language over time. In computational approaches, a possible merging of symbolic underspecification with neural inference is needed to allow for uncertainty and to influence multimodal grounding. Deployment scenarios should consider ethical issues and use the designs to responsibly reduce underdetermination. By following these principles, a design framework can be generated for AI systems so that they understand language similarly to how we do, respecting the flexibility of language and human-like vagueness.

CONCLUSION

Our thorough analysis of semantic underspecification has charted its development through five key stages:

1. Classical semantics, as discussed by Frege, Russell, and Davidson, aimed to create fully explicit and richly detailed meaning representations. However, they encountered some tricky challenges along the way, especially with indexicals, quantifier scope, and belief contexts.
2. Formal underspecification frameworks like QLF, MRS, NDCS, and Hole Semantics introduced computational flexibility, particularly for structural ambiguity, but they still did not fully capture real-world inference.
3. Cognitive-pragmatic views highlighted humans' preference for context-based interpretative simplicity: speakers often produce underspecified language, and listeners use inference to clarify meaning.
4. Hybrid approaches (e.g., Glue Semantics, UMR), combining symbolic and neuro-symbolic methods, maintain structural ambiguity while supporting partial inference—showing potential but lacking full uncertainty modeling.
5. Recent empirical research (Pezzelle 2023; Wildenburg et al. 2024) shows that even advanced language and multimodal models usually imperfectly detect

underspecification and interpret meaning in a deterministic way, lacking human-like awareness of uncertainty.

We have observed that while symbolic formalisms manage structural ambiguity, they are unable to resolve issues that are dependent on context. Neural language models are highly effective at detecting underspecification but tend to condense meanings into confident, single-point estimates, which conflicts with the expectation of ambiguity-aware interpretation, as noted by the ACL Anthology.

Although emerging entropy-based models elucidate the origins of grammatical underspecification in gender systems and other fields, Natural Language Processing (NLP) systems predominantly focus on English and tend to neglect typological diversity.

A significant number of evaluation datasets exclude genuinely ambiguous or underspecified data. Consequently, models may achieve satisfactory results on superficial metrics but encounter difficulties in real-world discourse, where underdetermination is commonplace.

Addressing semantic underspecification extends beyond purely academic pursuits; it represents a philosophical evolution within linguistic theory and presents a substantial technological obstacle in the development of resilient and interpretable language systems. Future models must carefully balance the implementation of formal frameworks that retain ambiguity, inference techniques that integrate contextual information, and deployment environments that recognise when information is insufficient to provide a definitive answer. Attaining this equilibrium is crucial for AI systems to replicate the adaptability and communicative accuracy inherent in human language.

REFERENCES

AlBzour, A. (2016). Cognitive Systematicity of Semantic Change: Cross-Linguistic Evidence. (2016). *International Journal of Applied Linguistics and English Literature*, 5(3). https://doi.org/10.7575/aiac.ijalel.v.5n.3p.91

Alshawi, H., & van Eijck, J. (1991). Levels of logical form in the CLE. In H. Alshawi (Ed.), *The Core Language Engine* (pp. 27–89). MIT Press.

Arnold, J. E. (2008). Reference Production: Production-Internal and Addressee-Oriented Processes. *Language and Cognitive Processes*, 23(4), 495–527. DOI:10.1080/01690960801920099

Arts, A., Maes, A., Noordman, L., & Jansen, C. (2011). Overspecification facilitates object identification. *Journal of Pragmatics*, 43(1), 361–374. https://doi.org/10.1016/j.pragma.2010.07.013

Bonami, O., & Godard, D. (2007). Parentheticals in Underspecified Semantics: The Case of Evaluative Adverbs. Research on Language and Computation, 5(4), 391–413. https://doi.org/10.1007/s11168-007-9037-z

Bos, J. (1995). *Predicate logic unplugged.* In Proceedings of the 10th Amsterdam Colloquium (pp. 133–142). ILLC.

Bos, J. (1997). *Predicate logic unplugged: Applications of logic to semantics.* PhD thesis, University of Amsterdam.

Davidson, D. (1967). Truth and Meaning. *Synthese*, *17*(3), 304–323. https://doi.org/10.1007/BF00485035

Degen, J. (2019). Modeling the Effect of Context on Vagueness. *Semantics and Linguistic Theory*, 29, 393–414. DOI:10.3765/salt.v29i0.4614

Devitt, M. (2021). Semantic polysemy and psycholinguistics. *Mind & Language*, 36(1), 134–157. https://doi.org/10.1111/mila.12327
Dziri, N., et al. (2022). Faithful Reasoning in LLMs. *NeurIPS*. DOI:10.48550/arXiv.2208.14271
E.G., M., & Opande, I. N. (2021). A Davidsonian Truth-theoretic Semantics Treatment of an EkeGusii Proverb. *Macrolinguistics,* 9(15), 68–94. https://doi.org/10.26478/ja2021.9.15.4
Egg, M. (2019). Semantic underspecification. In C. Maienborn, K. von Heusinger, & P. Portner (Eds.), *Semantics – Lexical Structures and Adjectives* (pp. 272–320). De Gruyter Mouton. https://doi.org/10.1515/9783110626391-009
Egg, M., Koller, A., & Niehren, J. (2001). The Constraint Language for Lambda Structures. *Journal of Logic, Language and Information,* 10(4), 457–485. https://doi.org/10.1023/a:1017964622902
Fecher, H., & Steffen, M. (2005). Characteristic μ-Calculus Formulas for Underspecified Transition Systems. *Electronic Notes in Theoretical Computer Science,* 128(2), 103–116. https://doi.org/10.1016/j.entcs.2004.11.031
Frank, M. C., & Goodman, N. D. (2012). Predicting Pragmatic Reasoning in Language Games. *Science*, 336(6084), 998. https://doi.org/10.1126/science.1218633
Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, 100, 25–50. https://doi.org/10.1515/9783110333200.25
Frisson, S. (2009). Semantic Underspecification in Language Processing. *Language and Linguistics Compass*, 3(1), 111–127. https://doi.org/10.1111/j.1749-818x.2008.00104.x
Frisson, S., & Pickering, M. J. (2001). Obtaining a Figurative Interpretation of a Word: Support for Underspecification. *Metaphor and Symbol*, 16(3–4), 149–171. https://doi.org/10.1080/10926488.2001.9678893
Futrell, R. & Levy, R. P., (2019) "Do RNNs learn human-like abstract word order preferences?", *Society for Computation in Linguistics* 2(1), 50-59. doi: https://doi.org/10.7275/jb34-9986
Gibson, E., et al. (2019). How Efficiency Shapes Human Language. *Trends in Cognitive Sciences*, 23(5), 389–407. DOI:10.1016/j.tics.2019.02.003
Gilbert, R. A., Davis, M. H., Gaskell, M. G., & Rodd, J. M. (2021). The relationship between sentence comprehension and lexical-semantic retuning. *Journal of Memory and Language*, 116, 104188. https://doi.org/10.1016/j.jml.2020.104188
Goodman, M. W. (2020). Penman: An open-source library and tool for AMR graphs. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations, pages 312–319, Online. Association for Computational Linguistics.
Griesel, H. (2007). Reform of the construction of the number system with reference to Gottlob Frege. *ZDM*, 39(1–2), 31–38. https://doi.org/10.1007/s11858-006-0003-2
Hedger, J. A. (2013). Meaning and racial slurs: Derogatory epithets and the semantics/pragmatics interface. Language & Communication, 33(3), 205–213. https://doi.org/10.1016/j.langcom.2013.04.004
Jackson, B. B. (2020). Model-theoretic semantics as model-based science. *Synthese*, 199(1–2), 3061–3081. https://doi.org/10.1007/s11229-020-02924-5
Koller, A., Niehren, J. & Striegnitz, K. (2000) Relaxing Underspecified Semantic Representations for Reinterpretation. *Grammars* 3, 217–241 https://doi.org/10.1023/A:1009948924590

KRIPKE, S. A. (2008). Frege's Theory of Sense and Reference: Some Exegetical Notes¹. *Theoria*, 74(3), 181–218. https://doi.org/10.1111/j.1755-2567.2008.00018.x

Lai, Y.-Y., & Piñango, M. M. (2017). *Comprehending underspecified meaning: resolving ambiguity by contextual and conceptual search*. Center for Open Science. https://doi.org/10.31234/osf.io/rxqz8

Le, N. B. (2024). The Semantic Analysis of TOWARDS from the Perspective of Cognitive Linguistics. *International Journal of Language Instruction*, 3(3), 22–35. https://doi.org/10.54855/ijli.24332

Lesmo, L., Robaldo, L., & Villata, S. (2007). *Meaning-Text Theory and Dependency Tree Semantics: An account of underspecification*. In Proceedings of the 3rd International Conference on the Meaning-Text Theory (MTT-2007) (pp. 249–258). Klagenfurt, Austria.

MacGregor, L. J., Rodd, J. M., Gilbert, R. A., Hauk, O., Sohoglu, E., & Davis, M. H. (2020). The Neural Time Course of Semantic Ambiguity Resolution in Speech Comprehension. *Journal of Cognitive Neuroscience*, 32(3), 403–425. https://doi.org/10.1162/jocn_a_01493

Manshadi, M., Gildea, D., & Allen, J. F. (2018). A Notion of Semantic Coherence for Underspecified Semantic Representation. *Computational Linguistics*, *44*(1), 39–83. https://doi.org/10.1162/coli_a_00307

Mao, R., He, K., Zhang, X., Chen, G., Ni, J., Yang, Z., & Cambria, E. (2023). A Survey on Semantic Processing Techniques. *ArXiv*. https://arxiv.org/abs/2310.18345

McElvenny, J. (2018). *The Meaning of Meaning.* In Edinburgh University Press. Edinburgh University Press. https://doi.org/10.3366/edinburgh/9781474425032.003.0002

Mirdehghan, M. (2008). Cognitive Linguistics from a Semantic Perspective. *Pazand Quarterly* 4 (12), 115-130, 2008.

Nicoletti, S. M., Peppelman, M., Kolb, C., & Stoelinga, M. (2023). Model-based joint analysis of safety and security:Survey and identification of gaps. *Computer Science Review*, 50, 100597. https://doi.org/10.1016/j.cosrev.2023.100597

Niehren, J., & Koller, A. (2001). *Dominance Constraints in Context Unification. In Lecture Notes in Computer Science* (pp. 199–218). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-45738-0_12

Oliveira, B. F. de. (2015). A FILOSOFIA SEMÂNTICA TRADICIONAL: Uma abordagem entre Frege e Russell. *Alamedas*, 3(2). https://doi.org/10.48075/ra.v3i2.12793

Panther, K.-U., & Thornburg, L. L. (2023). Hypallage is a rare bird. Not. *Lingua*, 289, 103511. https://doi.org/10.1016/j.lingua.2023.103511

Pezzelle, S. (2023). *Dealing with Semantic Underspecification in Multimodal NLP* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2306.05240

Piantadosi, S. T., et al. (2012). The Communicative Function of Ambiguity. *Cognition*, 122 (3), 280–291. DOI:10.1016/j.cognition.2011.10.004

Qadir Sabir, A. (2023). The Relationship between the World and Language in Bertrand Russell's Philosophy. *Journal of University of Raparin*, 10(4), 302–321. https://doi.org/10.26750/vol(10).no(4).paper14

Rasin, E., & Aravind, A. (2020). The nature of the semantic stimulus: the acquisition of every as a case study. *Natural Language Semantics*, 29(2), 339–375. https://doi.org/10.1007/s11050-020-09168-6

Russell, B. (1905). On Denoting. *Mind*, 14(56), 479–493. https://doi.org/10.1093/mind/XIV.4.479
Sag, I. A., & Wasow, T. (2014). Flexible Processing and the Design of Grammar. *Journal of Psycholinguistic Research*, 44(1), 47–63. https://doi.org/10.1007/s10936-014-9332-4
Sanford, A. J., & Graesser, A. C. (2006). Shallow Processing and Underspecification. *Discourse Processes*, 42(2), 99–108. https://doi.org/10.1207/s15326950dp4202_1
Sanford, A. J., & Sturt, P. (2002). Depth of processing in language comprehension: not noticing the evidence. *Trends in Cognitive Sciences*, 6(9), 382–386. https://doi.org/10.1016/s1364-6613(02)01958-7
Schwen Blackett, D., Varkey, J., Wilmskoetter, J., Roth, R., Andrews, K., Busby, N., Gleichgerrcht, E., Desai, R. H., Riccardi, N., Basilakos, A., Johnson, L. P., Kristinsson, S., Johnson, L., Rorden, C., Spell, L. A., Fridriksson, J., & Bonilha, L. (2022). Neural network bases of thematic semantic processing in language production. *Cortex*, 156, 126–143. https://doi.org/10.1016/j.cortex.2022.08.007
Soames, S. (2014). *Language, Meaning, and Information. In Analytic Philosophy in America.* Princeton University Press. https://doi.org/10.23943/princeton/9780691160726.003.0003
Stanley, J. (1996). Truth and Metatheory in Frege. *Pacific Philosophical Quarterly*, 77(1), 45–70. https://doi.org/10.1111/j.1468-0114.1996.tb00158.x
Su, C., Peng, Y., Huang, S., & Chen, Y. (2020). A Metaphor Comprehension Method Based on Culture-Related Hierarchical Semantic Model. *Neural Processing Letters*, 51(3), 2807–2826. https://doi.org/10.1007/s11063-020-10227-6
Suarez, A. G. (1996). Reference Without Sense: An Examination of Putnam's Semantic Theory. In *Philosophy and Cognitive Science: Categories, Consciousness, and Reasoning* (pp. 115–132). Springer Netherlands. https://doi.org/10.1007/978-94-015-8731-0_6
Van Gysel, J. E. L., Vigus, M., Chun, J., Lai, K., Moeller, S., Yao, J., O'Gorman, T. J., Cowell, A., Croft, W., Huang, C. R., Hajic, J., Martin, J. H., Oepen, S., Palmer, M., Pustejovsky, J., Vallejos, R., & Xue, N. (2021). Designing a Uniform Meaning Representation for Natural Language Processing. *Künstliche Intelligenz, 35*(3), 343–360.
Van Gysel, J., et al. (2024). Uniform Meaning Representation for Cross-Lingual Underspecification. *LREC-COLING*. DOI:10.48550/arXiv.2401.XXXXX
Vasile, M. D. (2013). The endless truth in the P. M. formal systems (In honor of the Principia Mathematica (1910-1913) of Bertrand Russell). Balkan Journal of Philosophy, 5(2), 139–146. https://doi.org/10.5840/bjp2013522
Wildenburg, F., Hanna, M., & Pezzelle, S. (2024). Do Pre-Trained Language Models Detect and Understand Semantic Underspecification? Ask the DUST! (Version 2). *arXiv*. https://doi.org/10.48550/ARXIV.2402.12486