



Semantic Structure of Indonesian Verbs *Mengira* and *Menyangka*: A Cognitive Explication Study

Afid Andono¹, Tri Mastoyo Jati Kesuma²
^{1,2}Magister Linguistik, Universitas Gadjah Mada

Article Info	Abstract
Received: 2026-02-08 Revised: 2026-04-06 Accepted: 2026-09-10 Keywords: <i>Cognitive explication;</i> <i>Epistemic verbs;</i> <i>Mengira;</i> <i>Menyangka;</i> <i>NSM</i> DOI: 10.24256/ideas.v14i2.9638 Corresponding Author: Afid Andono afidandono@mail.ugm.ac.id Magister Linguistik, Universitas Gadjah Mada	<i>Epistemic mental verbs are essential in signaling a speaker's degree of certainty, yet the semantic boundaries between the Indonesian verbs mengira and menyangka remain ambiguous due to lexical overlap and circular definitions in traditional dictionaries. This study aims to uncover the distinct semantic structures of these two near-synonyms by identifying their diagnostic features and cognitive profiles. Adopting a qualitative-descriptive design, the research utilizes a bottom-up approach that integrates empirical data from the Indonesian Leipzig Corpora Collection with the Natural Semantic Metalanguage (NSM) framework. The analysis of 400 purposive concordance lines reveals that mengira and menyangka occupy different cognitive spaces. The results indicate that mengira functions primarily as a verb of intellectual miscalculation, representing a tentative supposition that is frequently invalidated by factual correction, often collocated with ternyata. Conversely, menyangka operates as a verb of failed expectation or suspicion, encoding a clash between a pre-existing mental model and a contradictory reality, evidenced by its high frequency in negative constructions (tidak menyangka). The NSM explications succeed in bypassing circularity by using universal primes to map these distinct "cognitive scenarios." These findings prove that while both lexemes belong to the "thinking" category, mengira is more intellectual-proactive, whereas menyangka is more affective-reactive. This study provides a robust model for understanding the Indonesian mental lexicon and offers a practical solution to the synonymy problem in pedagogical contexts.</i>

1. Introduction

Mental verbs are essential in human communication because they show how speakers process and share their internal thoughts. Within this category, epistemic verbs function to signal a speaker's epistemic stance, which is the speaker's attitude or degree of certainty toward the truth of the information. This process is closely related to construal, or the subjective way a speaker chooses to perceive and describe a situation. These choices help

manage mental spaces, which are the "mental areas" where a speaker organizes their information and beliefs.

In Indonesian, this cognitive process is clearly seen in the verbs *mengira* and *menyangka*. Both are frequently used to express a "supposition" or "tentative belief," yet their usage often overlaps, creating a semantic blur. While both verbs belong to the broader category of "thinking," they diverge in their relation to "factivity", the degree to which a speaker commits to the truth of a claim (Lyons, 1995). Lexical ambiguity in mental verbs often stems from the complex mapping between linguistic forms and the conceptual "mental spaces" they inhabit (Evans, 2019). Without a precise semantic decomposition, these verbs remain indistinguishable in pedagogical contexts, despite their different "cognitive profiles."

Current scholarship in lexical semantics has increasingly moved toward compositional models to resolve such ambiguities. Wierzbicka (1996) and Goddard (2010, 2021) have established that mental lexicon analysis requires a universal metalanguage to remain objective and cross-linguistically valid. Previous studies on Indonesian mental verbs, such as those by Mulyadi (2014, 2016), have successfully mapped the semantic structures of verbs like *tahu* (know) and *percaya* (believe) using this framework. Furthermore, Wijayanti & Sudarmini (2022) explored cognitive features in Indonesian verbs but focused on broad classifications rather than the microscopic differentiation of near-synonyms. Other research has utilized systemic functional linguistics to view mental processes as grammatical metaphors (Sembiring et al., 2021), yet the underlying "cognitive scenario" of suppositional verbs like *mengira* remains under-researched, particularly in how they manage the speaker's "epistemic stance" the interactional display of knowledge and certainty (Heritage, 2012).

Despite the aforementioned progress, a comparative study specifically targeting *mengira* and *menyangka* is notably absent. Most existing literature treats them as interchangeable members of the "thinking" category, neglecting the fact that near-synonyms often exhibit different distributional patterns that reflect distinct conceptualizations (Geeraerts, 2010). Furthermore, there is a lack of research that combines large-scale empirical data from the Indonesian Leipzig Corpora Collection with the rigorous reductive explication of Natural Semantic Metalanguage (NSM). Without a corpus-based approach, semantic analysis risks relying solely on the researcher's intuition, which may miss the nuances of frequency and collocation that signal cognitive differences. This lack of empirical grounding often leads to "armchair linguistics," which fails to account for the dynamic nature of meaning in actual usage (Sinclair, 1991).

This study addresses these gaps by posing two primary questions: (1) what are the diagnostic features that differentiate *mengira* and *menyangka* in natural discourse, and (2) how can their semantic structures be explicated using NSM primes? The objective is to provide a precise cognitive map for these two verbs, elucidating how they encode different levels of expectation and factual realization. The novelty of this research lies in its "bottom-up" methodology using corpus-driven evidence to build "top-down" cognitive explications thereby bridging the gap between empirical frequency and theoretical semantics. By synthesizing data-driven insights with the NSM framework, this study offers a more robust and objective model for understanding the Indonesian mental lexicon, ensuring that the resulting explications are both cognitively plausible and empirically verifiable.

2. Method

This section delineates the systematic procedures and theoretical frameworks employed to investigate the semantic differentiation between *mengira* and *menyangka*. The methodology is constructed to ensure empirical grounding by synthesizing large-scale corpus data with a rigorous qualitative analysis. By adopting a structured approach that

moves from linguistic evidence to conceptual explication, this research aims to provide an objective map of the Indonesian mental lexicon while addressing the limitations of traditional lexicography. The following sub-sections detail the specific research design, data sources, and the analytical stages involved in this study.

Research Design

This research employs a qualitative-descriptive design rooted in the intersecting paradigms of cognitive linguistics and lexical semantics. The study is strategically designed to utilize a "bottom-up" approach, facilitating a transition from granular empirical evidence to high-level theoretical abstraction. This design ensures that the study provides an authentic "cognitive portrait" of the Indonesian speaker's mental state as encoded in *mengira* and *menyangka*. By prioritizing actual communicative events over mere researcher intuition, the study adheres to the principle that meaning is inherently grounded in usage, as posited by Langacker (2008) regarding the subjective nature of linguistic construal.

Data Source

The subjects of this research are the linguistic tokens of *mengira* and *menyangka* captured within naturalistic discourse. The data were elicited from the Indonesian Leipzig Corpora Collection, utilizing the "ind_mixed_2012" and "ind_news_2023" datasets to ensure a representative balance between formal journalistic registers and diverse informal contexts.

Data Collection

Data were collected through a computerized search, yielding a purposive sample of 200 concordance lines for each verb (400 tokens in total). In this study, "high-quality concordance lines" are defined as sentences that provide a complete and unambiguous semantic context. Specifically, the selection prioritized tokens situated within rich syntagmatic environments, such as those containing explicit complement clauses (e.g., *bahwa...*) or contrastive markers (e.g., *ternyata...*) that clearly signal the speaker's epistemic positioning. To ensure a comprehensive analysis of the "cognitive arc" of each utterance, the search was expanded to include two to three sentences surrounding each target token, following the principles of extended context in corpus linguistics (Sinclair, 1991).

Data Analysis

The data analysis follows a systematic four-stage reductive process within the Natural Semantic Metalanguage (NSM) framework:

Contextual identification: Categorizing the collocations and syntactic patterns of the sampled tokens.

Componential analysis: Extracting diagnostic features such as [$\pm$ certainty] and [$\pm$ expectation].

Reductive explication: Formulating semantic explications using universal, indefinable primes such as THINK, KNOW, TRUE, and NOT.

Cognitive interpretation: Synthesizing these structures to provide a robust model of the Indonesian mental lexicon (Heritage, 2012).

The semantic analysis is conducted using the Natural Semantic Metalanguage (NSM) framework, based on the premise that all complex meanings can be decomposed into a set of 65 universal and indefinable conceptual atoms known as "semantic primes". These primes represent the fundamental building blocks of human thought and are cross-linguistically valid (Wierzbicka, 1996; Goddard, 2021).

Table 1. Inventory of Semantic Primes (Abridged)

Category	Semantic Primes
Substantives	I, YOU, SOMEONE, SOMETHING, PEOPLE, BODY
Mental predicates	THINK, KNOW, WANT, FEEL, SEE, HEAR
Speech	SAY, WORDS, TRUE
Actions, events, movement	DO, HAPPEN, MOVE
Existence and possession	THERE IS, HAVE
Life and death	LIVE, DIE
Descriptors	GOOD, BAD, BIG, SMALL
Time	WHEN, NOW, AFTER, BEFORE, A LONG TIME, A SHORT TIME
Space	WHERE, HERE, ABOVE, BELOW, FAR, NEAR, SIDE, INSIDE
Logical concepts	NOT, MAYBE, CAN, BECAUSE, IF
Augmentors and Quantifiers	VERY, MORE, ALL, SOME, ONE, TWO

Research Limitation

A notable limitation of this study is its primary reliance on written data extracted from the Indonesian Leipzig Corpora Collection, which consists largely of news articles and formal mixed texts. Consequently, the semantic nuances identified in *mengira* and *menyangka* specifically reflect the patterns of formal or written Indonesian. These findings may exhibit variations in spontaneous spoken discourse or specific regional dialects, where prosodic features and informal social contexts might further influence the speaker's epistemic weight. Despite these constraints, the study provides a significant advancement in the objective mapping of the Indonesian mental lexicon.

3. Results

The results of the analysis indicate that while *mengira* and *menyangka* are often used interchangeably in casual discourse, they occupy distinct cognitive spaces in the Indonesian mental lexicon. The following subheadings detail the findings regarding their syntactic patterns, diagnostic features, and formal NSM explications.

Semantic Structure of *Mengira*

The analysis of the Indonesian Leipzig Corpora Collection reveals that the verb *mengira* is predominantly used to express a tentative epistemic stance that is subsequently invalidated by new information. Syntactically, *mengira* frequently appears in the pattern [**S + *mengira* + *bahwa* + Clause**] or [**S + *mengira* + NP**]. A salient feature of *mengira* is its strong association with the conjunction *ternyata* (it turns out), which functions as a marker of factual correction. For instance, corpus data shows sentences such as: "*Awalnya saya mengira dia adalah seorang guru, tetapi ternyata dia seorang dosen*" (Initially, I thought he was a teacher, but it turns out he is a lecturer). In this context, *mengira* represents a "supposition of truth" held by the speaker at a specific time (T₁), which is later recognized as an "error of fact" at T₂.

The primary semantic component of *mengira* is its tentative-erroneous nature. Unlike the verb *tahu* (know), which implies a [factive] component where the proposition is true, *mengira* implies that the speaker thought something was true because they did not have enough information to know otherwise. This creates a "subjective belief space" that is

fragile and prone to revision. The epistemic certainty of *mengira* is relatively low, as it lacks the evidentiary backing of *yakin* (certain) or the deep internal commitment of *percaya* (believe).

The cognitive scenario of *mengira* can be formally explicated using NSM semantic primes as follows:

Explication of *mengira*:

someone X thinks like this at some time:

"I think something like this: this is true"

this someone thinks this because this someone does not know everything about this something at a later time, this someone can know like this: "this is not true"

it is like this because this someone thought something before that was not true

Semantic Structure of *Menyangka*

In contrast, *menyangka* carries significant weight regarding expectation and suspicion, appearing frequently in negative constructions (*tidak menyangka*). It typically occurs when an event contradicts the speaker's prior mental model. An example from the dataset is: "*Penduduk desa tidak menyangka bahwa banjir akan datang secepat itu*" (The villagers did not expect/think that the flood would come that fast). Here, *menyangka* does not just mean "thinking," but "having an expectation about what could happen."

The diagnostic feature of *menyangka* is its [$\pm$ expectation] component. While affirmative usage can border on "suspicion" (*kecurigaan*), its primary function is to mark a clash between expectations and reality. However, its most common cognitive function in modern Indonesian is to mark a **clash between expectation and reality**. This makes *menyangka* more "reactive" than *mengira*, which is more "proactive" in forming initial guesses.

The cognitive structure of *menyangka* is explicated in the NSM framework as follows:

Explication of *menyangka*:

someone X thinks like this at some time:

"I think: something like this could happen"

"I think: I know what this something is like"

because of this, this someone thinks: "this is how it is"

at another time, something happens; this someone sees it

because of this, this someone thinks: "this is not what I thought before"

this someone feels something because of this

Comparative Analysis of Diagnostic Features

To provide a concise summary of the findings, the following table illustrates the differentiation between the two verbs based on the epistemic features identified during the analysis of the Leipzig Corpora.

Table 2. Comparative Analysis of Diagnostic Features

Feature	Mengira	Menyangka
Epistemic Certainty	Low (Tentative)	Moderate (Based on Expectation)
Factual Correction	High (Often followed by <i>ternyata</i>)	High (Often followed by surprise/reaction)
Primary Nuance	Miscalculation/Erroneous guess	Expectation/Suspicion/Anticipation
Negative Polarity	Rare (<i>tidak mengira</i> is less common)	High (<i>tidak menyangka</i> is highly frequent)

Feature	Mengira	Menyangka
Core NSM Prime	THINK (this is true)	THINK (something can happen/is like this)

The table suggests that *mengira* is more "intellectual" in its miscalculation, focusing on the error of the proposition itself. *Menyangka*, however, is more "affective-cognitive," focusing on the failure of the speaker's expectation or mental preparedness for a specific reality. This distinction proves that although both are categorized as "mental verbs of thinking," their underlying cognitive scenarios reflect different ways in which Indonesian penutur (speakers) process information and surprise.

Mengira focuses on the intellectual error of a proposition, while *menyangka* focuses on the affective-cognitive failure of a speaker's preparedness for reality.

4. Discussion

The comparative analysis of *mengira* and *menyangka* reveals a fundamental divergence in their epistemic orientation, moving beyond the singular "supposition" category found in traditional dictionaries. This section synthesizes these patterns into broader cognitive and pedagogical implications.

Epistemic Continuum and Cognitive Positioning

The findings suggest that *mengira* and *menyangka* occupy distinct niches within the Indonesian epistemic continuum, located between certainty (*yakin*) and ignorance (*tidak tahu*). *Mengira* aligns with a neutral "guess," where the speaker acknowledges a lack of information from the outset. In contrast, *menyangka* is deeply rooted in "predictive processing," where a pre-existing mental model is held with subjective confidence until disrupted by reality.

The ubiquitous use of *tidak menyangka* in the corpus indicates that this verb manages the failure of anticipation rather than a mere lack of thought. While *mengira* represents a proactive intellectual attempt to assign truth, *menyangka* serves as a reactive marker for when the world does not match the speaker's internal expectations.

Critical Perspective and Alternative Interpretations

While the NSM framework effectively bypasses circularity by using universal primes, it is important to acknowledge its inherent limitations. The reductive nature of NSM may occasionally oversimplify the affective nuances of mental verbs. For instance, the "suspicion" aspect of *menyangka* in affirmative contexts might be influenced by social hierarchies and pragmatic politeness factors that are difficult to capture solely through semantic atoms.

Alternative interpretations, such as those from Systemic Functional Linguistics, might view these verbs as grammatical metaphors for interpersonal alignment rather than purely cognitive processes. However, the cognitive approach used here provides a more granular look at the individual speaker's internal "mental area" management.

NSM Implications and Theoretical Synthesis

The use of semantic primes to explicate these verbs resolves the circularity found in conventional definitions. By reducing *mengira* to the scenario of "thinking something is true because of not knowing everything," we align this research with Goddard's (2010) findings on the conceptual semantics of 'knowing'. The explication for *mengira* captures the "fragility" of the belief, which explains its frequent collocation with *ternyata*. This collocation acts as a cognitive "reset" button, where the speaker transitions from a state of *THINK* to a state of *KNOW* (*this is not true*).

For *menyangka*, the integration of the component "something like this could happen" aligns with the cognitive linguistics' view of "mental imagery" and "construal" (Langacker, 2008). The semantic structure of *menyangka* is more complex than *mengira* because it involves an interaction between the speaker's past expectations and the current perceptual input. As noted by Wierzbicka (1996), mental verbs often encode cultural values; in Indonesian, the nuances of *menyangka* which can shift from "expectation" to "suspicion" (*kecurigaan*) in affirmative contexts reflect a social sensitivity toward the hidden intentions of others. This "suspicion" aspect is a unique pragmatic byproduct of the [THINK: I know what this someone is like] component identified in the NSM explication.

Contribution to Knowledge and Pedagogical Impact

This study explicitly adds new knowledge to the Indonesian mental lexicon by providing a verifiable, non-circular model of suppositional verbs. It resolves the long-standing "synonymy problem" in Indonesian lexicography by identifying the [$\pm$ expectation] component as the primary diagnostic differentiator.

For Indonesian for Foreign Learners (BIPA), these findings offer a practical pedagogical tool. Instead of teaching these verbs as interchangeable, instructors can present *mengira* as a tool for intellectual hypotheses and *menyangka* for experiential expectations. This data-driven approach ensures that learners develop a native-like intuition for managing epistemic responsibility in interaction.

Pragmatic and Pedagogical Consequences

The differentiation established in this study has profound implications for both pragmatic theory and the teaching of Indonesian as a Foreign Language (BIPA). Pragmatically, the choice between *mengira* and *menyangka* serves as an "epistemic stance" marker (Heritage, 2012). A speaker uses *mengira* to downplay their epistemic responsibility for an error ("*I just guessed wrong*"), whereas *menyangka* is used to emphasize the unexpected nature of a situation ("*I had a different expectation*").

In a pedagogical context, the current findings provide a solution to the "synonymy problem" that often confuses advanced learners. Instead of teaching these verbs as synonyms, instructors can present them as cognitive tools: *mengira* for intellectual hypotheses and *menyangka* for experiential expectations. This approach, grounded in the systematic analysis of the Leipzig Corpora, ensures that learners develop a "native-like" intuition for the subtle epistemic nuances that define the Indonesian language.

5. Conclusion

This study explicitly addresses and resolves the long-standing problem of semantic circularity in Indonesian lexicography regarding mental verbs. By providing a non-circular, empirical model based on Natural Semantic Metalanguage (NSM), this research offers a precise tool for differentiating near-synonyms that traditional dictionaries often conflate. These findings provide a new foundation for both theoretical Indonesian semantics and practical pedagogical applications in Indonesian for Foreign Learners (BIPA) contexts.

The research concludes that *mengira* and *menyangka* represent distinct cognitive profiles within the Indonesian epistemic continuum. *Mengira* is characterized as a verb of intellectual miscalculation, where a tentative hypothesis is frequently invalidated by factual correction. In contrast, *menyangka* functions as a verb of failed expectation or suspicion, encoding a clash between a pre-existing mental model and a contradictory reality. The resulting NSM explications successfully bypass circularity, providing a verifiable model of how these verbs manage the speaker's epistemic stance.

However, certain limitations must be acknowledged. The findings are primarily based on written data from news articles and formal texts in the Leipzig Corpora. Consequently, these semantic nuances may vary in spontaneous spoken discourse, where prosodic

features and informal social contexts further influence epistemic weight. Furthermore, this study focuses on only two specific lexemes within the vast category of Indonesian mental verbs.

Based on these findings, several recommendations are offered for future inquiry. First, further studies should expand the analysis to include a wider array of suppositional verbs, such as *menduga* and *menaksir*, to create a comprehensive map of the Indonesian epistemic continuum. Second, researchers are encouraged to apply interactional linguistics to examine these verbs in spoken discourse. Finally, BIPA instructional materials should incorporate these cognitive-semantic distinctions to foster a "native-like" understanding of epistemic tools in Indonesian communication.

6. Acknowledgement

The author would like to express his deepest gratitude to several parties whose support was instrumental in the completion of this research. First and foremost, a special word of thanks is directed to Dr. Yohanes Tri Mastoyo Jati Kesuma, M.Hum., as the research supervisor, for his invaluable guidance, meticulous suggestions, and constant encouragement throughout the development of this study. His profound expertise in Indonesian linguistics and methodology significantly improved the theoretical depth and analytical rigor of this article.

Furthermore, the author gratefully acknowledges the financial support provided by the Indonesia Endowment Fund for Education (LPDP) through the Magister Scholarship program. This research would not have been possible without the funding and academic opportunities facilitated by this institution. Finally, appreciation is also extended to the Faculty of Cultural Sciences, Universitas Gadjah Mada, for providing the research facilities and academic environment that fostered the intellectual growth necessary for this work.

7. References

- Evans, V. (2019). *Cognitive linguistics: A complete guide*. Edinburgh: Edinburgh University Press.
- Geeraerts, D. (2010). *Theories of Lexical Semantics*. Oxford: Oxford University Press.
- Goddard, C. (2010). *Semantic Analysis: A Practical Introduction*. Oxford: Oxford University Press.
- Goddard, C. (2021). *Ten lectures on Natural Semantic Metalanguage: Exploring language, thought and culture*. Leiden: Brill.
- Goddard, C., & Wierzbicka, A. (2014). *Words and meanings: Lexical semantics across domains, languages, and cultures*. Oxford: Oxford University Press.
- Heritage, J. (2012). Epistemics in action: Action formation and terrestrial structures in social interaction. *Research on Language and Social Interaction*, 45(1), 1–29. <https://doi.org/10.1080/08351813.2012.646684>
- Langacker, R. W. (2008). *Cognitive grammar: A basic introduction*. Oxford: Oxford University Press.
- Lyons, J. (1995). *Linguistic semantics: An introduction*. Cambridge: Cambridge University Press.
- Mulyadi. (2014). *Struktur semantis verba mental bahasa Indonesia* (Unpublished doctoral dissertation). Universitas Sumatera Utara, Medan.
- Mulyadi. (2016). *Metabahasa semantik alami: Teori dan kajian empiris*. Medan: USU Press.
- Sembiring, B., Khorina, M., & Nugraha, D. G. (2021). Mental verbs in marketing texts: Systemic functional linguistics perspectives. *Linguistics and English Language Teaching Journal*, 9(1), 1–9.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Wierzbicka, A. (1996). *Semantics: Primes and universals*. Oxford: Oxford University Press.

Wijayanti, A., & Sudarmini, S. (2022). Analisis Kata Kerja Mental dalam Berita Utama Surat Kabar Kedaulatan Rakyat Periode Januari—Februari 2022. *Jurnal Serunai Bahasa Indonesia*, 19(2), 61–70.
<https://ejournal.stkipbudidaya.ac.id/index.php/je/article/view/680/429>