Misunderstanding the Concrete, Disagreeing about the Abstract: A Closer Look at Word Meaning Negotiation Triggers

Bill Noble[†] bill.noble@gu.se
Staffan Larsson[†] staffan.larsson@ling.gu.se
Jenny Myrendal[‡] jenny.myrendal@gu.se

[†]Dept. of Philosophy, Linguistics and Theory of Science
[‡]Department of Education, Communication and Learning
University of Gothenburg

Abstract

Word Meaning Negotiation (WMN) occurs when speakers explicitly address the meaning of a word or phrase – a *trigger expression* – often in response to either non-understanding or disagreement. This paper examines the lexical and semantic features of trigger expressions in a set of 393 WMNs from the NeWMe Corpus, the first large-scale annotated dataset of WMNs across spoken and online interactions. We analyze triggers by concreteness, sentiment, part of speech, interaction modality and form, distinguishing patterns between disagreement- and non-understanding-driven cases. The results shed light on how different kinds of expressions are likely to trigger different kinds of negotiations of meaning in dialogue. One observation is that abstract expressions are associated with disagreement about word meaning, while concrete expressions are relatively more associated with negotiations due to misunderstanding.

1 Introduction

In everyday interaction, both spoken and written, participants sometimes encounter moments in which the meaning of a word becomes problematic or contested. When participants notice that a word's meaning requires clarification — either due to a lack of understanding or because they challenge how it is being used — they may initiate a Word Meaning Negotiation (WMN): a sequence in which the meaning of a term becomes the explicit topic of discussion, prompting a shift from content-level talk to meta-linguistic engagement. This shift may be triggered by a clarification request (e.g., "What do you mean by...") or by an objection to the use of a particular term (e.g., "That is not (the meaning of)..."). WMNs unfold as interactional sequences in which participants seek to clarify, redefine, or contest word meaning through strategies such as explicification, exemplification, and contrasting (Myrendal, 2015, 2019).

Although WMNs are often used to resolve misunderstandings, they also serve rhetorical purposes by enabling participants to challenge or defend positions in argumentative discourse. This is particularly evident in discussions involving moral or ideological disagreement, where negotiating the meaning of a term can function as a strategic move to redefine the issue or shift the normative ground of the debate (Myrendal and Larsson, 2025; Larsson and Myrendal, 2024).

While previous work has focused on the interactional structure and functions of WMNs, less is known about what kind of words WMNs are *about*. This work investigates WMN from the perspective of the lexical items that trigger the negotiation. Are there differences in the trigger word features for WMNs initiated by a clarification (i.e., an expression of non-understanding; NONs) versus an expression of disagreement (DINs)? Do features of the trigger word predict the *scope* of the WMN; that is, whether the WMN concerns the word's *situated* meaning (how it is used in a particular utterance or discourse), or the word's *meaning potential* more broadly (or both)? In particular, this study aims to explore the lexical dimensions of WMN trigger words by asking the following research questions:

Research questions

RQ1 What trigger word features are predictive of the type of WMN (NON or DIN)?

RQ2 What trigger word features are predictive of the kind of meaning (*situated meaning* or *meaning potential*) that is the focus of a WMN?

To address these questions, we analyze 393 annotated WMNs from the NeWMe Corpus - the first large-scale dataset of Word Meaning Negotiations across both spoken and online interaction.

Proceedings of the 29th Workshop on the Semantics and Pragmatics of Dialogue, September 3–5, 2025, Bielefeld, Germany.

2 Background

WMNs are structured sequences in which interlocutors explicitly negotiate the meaning of a word or phrase, typically following a three-part pattern: a Trigger (the initial use of a potentially problematic word), an Indicator (a subsequent utterance that highlights or challenges the meaning of that word), and one or more Response turns that engage in meta-linguistic elaboration. This T-I-R (Trigger–Indicator–Response) structure is inspired by Varonis and Gass (1985)'s model of negotiated meaning, which also emphasizes the role of an initial problematic utterance, a signal of difficulty, and negotiated responses in second language interaction.

WMNs can be triggered by non-understanding (NON) or by disagreement (DIN), and they may concern a word's meaning in the specific context (situated meaning) or in general (potential meaning) (Myrendal, 2015; Norén and Linell, 2007). Here is an example of WMN Caused by Non-Understanding of Word Meaning (NON):

S1: I'm going to the doctor to get a full body scan tomorrow.
S2: What do you mean by full body scan?
S1: I mean a kind of X-ray where they can see all of the inflamed parts.

This example, taken from Myrendal (2015), illustrates a scenario where S1's use of the term "full body scan" serves as the trigger, introducing a word which is not fully understandable to S2. S2 then produces an indicator, explicitly requesting clarification about the meaning of "full body scan," making this phrase the trigger. In response, S1 provides an explanation, elaborating on the word to address the lack of understanding. This sequence demonstrates how WMNs initiated by non-understanding (NONs) focus on clarifying the meaning of specific terms to maintain mutual understanding in the conversation.

Next is an example of WMN Caused by Disagreement about Word Meaning (DIN):

S1: Telling children about Santa Claus is straight up lying to them.
S2: That's not what lying means at all!
S1: Of course it is, lying means not telling the truth and everyone knows Santa doesn't exist.

This example, drawn from Norén and Linell (2007), illustrates a WMN caused by disagreement about word meaning (DIN), where the focus shifts to negotiating differing perspectives on the meaning of a word. Here, S1's initial statement introduces the word "lying," which serves as the trigger. S2 challenges this usage by providing an indicator, asserting that the term "lying" does not apply in the given context and objecting to its use. In response, S1 elaborates on their understanding of the word, reinforcing their interpretation and connecting it to the situation at hand.

According to Noren and Linell (2005), words have *meaning potentials*, flexible semantic resources that can be activated and elaborated in various ways depending on the interactional context. A word's *situated meaning* is its meaning in a particular context of use.[1] In WMNs, participants collaboratively shape which aspects of a word's meaning potential are made relevant in the interactional context. Rather than aiming for a single correct or fixed interpretation, the negotiation centers on selecting and articulating interpretations that are contextually appropriate, socially acceptable, or strategically advantageous. What is at stake, then, is not an objective understanding of the term, but the interactive process of managing its semantic flexibility to achieve mutual intelligibility or advance particular stances. WMNs can focus on the trigger word's situated meaning by addressing what was meant by a particular speaker in a particular context of use; they can focus on meaning potential by more abstractly engaging what the word *can* mean; or they can include both kinds of meaning.

As noted in Gari Soler et al. (forthcoming), DINs tend to involve longer exchanges than NONs, averaging 7.2 turns compared to 3.5 turns. DINs display much greater variability in length. In our corpus, the longest NON contains 27 turns in total, while the longest DIN spans 268 turns. This highlights the more elaborate and prolonged nature of DINs, where participants engage in extended exchanges to explore and debate different interpretations of word meaning.

Previous research on WMNs has largely focused

[1] In multimodal computational linguistics, *situated meaning* sometimes refers more narrowly to the meaning of an expression in a particular shared perceptual context (e.g., Pustejovsky and Krishnaswamy (2020)). Here we use a broader notion of the term which includes social, conversational, and other aspects of context. For further discussion see §3 of Norén and Linell (2007).

on their sequential structure and interactional functions (Myrendal, 2015, 2019, 2025; Myrendal and Larsson, 2025). These studies have shown how speakers engage in strategies such as explicification, exemplification, and contrasting to address misunderstandings or disagreements about word use.

However, relatively little attention has been given to the lexical and semantic properties of the trigger expressions themselves. In particular, we lack systematic knowledge about whether certain word types - e.g., abstract vs. concrete, single-word vs. compound expressions, spoken vs. online interaction contexts, positive vs. negative sentiment, or natural kind vs. artefact nouns - are more likely to prompt negotiation. An exception is Garí Soler et al. (2023), who propose computational measures of lexico-semantic alignment in debates using contextualized word representations. Their findings show that shared lexical items do not necessarily imply shared semantic usage, suggesting a need for more fine-grained analysis of the expressions that become sites of explicit negotiation. This observation aligns with the broader view that meaning in interaction is not just a function of lexical semantics but of situated and strategic use. Our study addresses this empirical and conceptual gap by analyzing the lexical features of trigger expressions in the NeWMe Corpus.

3 Data

The primary data for this study comes from the NeWMe corupus (Section 3.1). We augment the WMNs from NeWMe with lexical semantic features of the trigger expression (concreteness and sentiment), which are drawn from other sources (Sections 3.2, 3.3).

3.1 The NeWMe corpus

The NeWMe[2] Corpus (Garí Soler et al., 2025) is the first large-scale annotated corpus of WMNs, encompassing spoken interactions sourced from the British National Corpus (BNC), Switchboard, and online discussions from Reddit's ChangeMyView forum. It includes annotations for WMN type (NON, DIN or Other[3]), focus (potential/situated/both), and spans for trigger words or expressions, indicator phrases, and negotiation spans[4].

The corpus contains 392 WMN instances. Each WMN includes an identified *trigger expression* — the word or phrase that is the focus of negotiation. Each WMN is also annotated with respect to meaning aspect (potential vs. situated). Furthermore, the NeWMe corpus specifies which source corpus (BNC, Switchboard or Reddit) each WMN comes from.

The distribution of the 392 WMNs according to Type of Word Meaning Negotiation is shown in Table 1. NONs are slightly more common than DINs overall.

Type	**#**	**%**
NON	216	55%
DIN	157	40%
Other	19	5%
Total	**392**	**100%**

Table 1: Distribution of NON, DIN, and Other types in the NeWMe corpus

The NeWMe data represents both online and spoken (and transcribed) interactions. The spoken WMNs originate from Switchboard and BNC. We refer to this parameter as Interaction Type. Differences depending on this parameter may be due to the medium of interaction (spoken vs. online written) but may also be due to the online data being mostly debates whereas the spoken data is more mixed with respect to dialogue genre. Interaction Type is distributed as shown in Table 2. This reflects the composition of the NeWMe corpus but note that it does not say how common WMNs are (e.g. in relation to the respective total number of lexical tokens) in the Online, Spoken (BNC) and Spoken (SW) corpora. We leave further investigation of this for future work.

In the NeWMe corpus, meaning aspect is distributed as shown in Table 3. About half of the WMNs concerned situated meanings, and about 1/4 concerned meaning potentials, with the remaining

[2] **Ne**gotiation of **W**ord **Me**aning

[3] The "Other" label is used for cases where word meaning was discussed without non-understanding or disagreement. These typically involve situations where one dialogue participant asks about a word and/or suggests an alternative word, which the other participant then confirms as appropriate.

[4] Inter annotator agreement results reported in Gari Soler et al. (forthcoming) show that inter-annotator agreement for WNM type was generally moderate, with agreement higher for NONs than for DINs. However, agreement on focus was lower, underscoring its subjective nature.

Interaction Type	#	%
Online	216	55%
Spoken (total)	176	45%
Spoken (BNC)	141	36%
Spoken (SW)	35	9%
Total	**392**	**100%**

Table 2: Distribution of interaction types in the sample

1/4 concerning both situated and potential meanings. See Figure 1 for examples of WMNs concerning situated meaning versus meaning potential.

Aspect	#	%
Situated	209	53%
Both	92	23%
Potential	91	23%
Total	**392**	**100%**

Table 3: Distribution of aspect of meaning

3.2 Concreteness Classification

To analyze the concreteness of lexical items, we relied on the concreteness ratings by (Brysbaert et al., 2014), who provide mean concreteness values for over 39,000 English word lemmas and common two-word expressions. The ratings, based on crowd-sourced judgments from more than 4,000 participants, use a five-point scale ranging from 1 (very abstract) to 5 (very concrete).

If such phrases were directly present in the Brysbaert dataset, we used the published rating. However, when a multi-word expression was not included in the concreteness or sentiment dataset, we instead fall back on using the head lemma of the multi-word expression[5] .

For instance, the expression *absolute power* is not contained in the abstractness or sentiment datasets, so the scores for the head word, *power* were used (mean concreneness = 1.93). Since *power* is rated as abstract, absolute power was also treated as abstract. Conversely, in *pop up tents*, the head noun *tent* (mean rating = 4.71) led us to classify the compound as concrete.

This head-based approach allowed us to systematically classify multi-word expressions while maintaining alignment with the theoretical understanding of concreteness as grounded in perceptual experience. Additional examples include human emotion, classified as abstract based on the head emotion (1.85), and smoke alarm (classified as concrete based on alarm, 4.36). When modifiers added evaluative or moral content (e.g., moral right, just war), we continued to prioritize the head noun (right, war) in line with syntactic structure, although we acknowledge that such modifiers can subtly influence perceived concreteness.

3.3 Lexical Sentiment Metrics

As a measure of the lexical sentiment of trigger words, we use SentiWordNet (version 3.0 Baccianella et al., 2010), which provides three sentiment-related metrics, measuring the *positivitiy*, *negativity* or *objectivity* (sentiment neutrality) of a lexical item. The dataset is constructed such that the three terms alwas sum to 1. As such, we employ only the positivity and negativity metrics (**PosSenti**, **NegSenti**) in our statistical model.

3.4 Other variables

In order to conduct this analysis, one of the authors annotated the trigger expression of each of the WMNs as belonging to one of five parts of speech: *noun*, *adjective*, *verb*, *adverb*, and *acronym*. Annotation occasionally involved inspecting the relevant interaction in the NeWMe corpus.

While most triggers consisted of single words, a substantial number of them were multi-word expressions (e.g., absolute power, pop up tents, moral right). Lexical form was manually annotated by one of the authors.

4 Descriptive statistics

We analyzed 393 WMN sequences from the NeWMe Corpus. In addition to the existing categorizations of WMNs described above, we categorized each trigger according to the following dimensions:

- Concreteness (abstract, concrete, or mixed)
- Sentiment (positive, negative, both, neither)
- Part of Speech
- Lexical Form (single-word, compound/multi-word phrase)

In WMNs where one form of the word is used in the trigger, but another form is used in the indicator, the trigger form has been chosen. Ie someone says

[5]To intedify head words, we use the SpaCy dependency parser with the `en_core_web_sm` model (version 3.8.0).

situated meaning	meaning potential
noun: *bell* | J9P/J9P_760	**noun:** *invisible fencing* | 4179-0/4179-0_4179-74/
A: Oh no not an alarm it's it's be too expensive, no just an internal **bell** to frighten the hell out of them. [...]	**A:** I see other people out there and they hit their dogs and try to— and those horrible collars that they put on them with invinc— **invisible fencing**, least I—
B: So what do you mean by a **bell**, [UNCLEAR] trying to visualize what you mean	**B:** Invisible what?
A: Yes. Well like [UNCLEAR] I mean we all know what a **bell** is, a **bell** which is set off by— by a human body coming in.	**A:** **Invisible fencing**, have you heard of that?
	B: No, what is that?
	A: It's— uh, it's a system you can put in your yard where you bury these little uh, transducers or emitters in your yard—
adjective: *recent* | FME/FME_18	**adjective:** *anthropogenic* | F8E/F8E_21
A: Is that is that **recent** or is that the old stuff the Venn diagrams?	**A:** Okay ? So we know that so far about fifty percent of our **anthropogenic** C O two has been locked away
B: What do you mean by **recent**?	**B:** What does **anthropogenic** mean?
A: Have you done it in the last sort of few weeks?	**A:** From human sources . For example can we continue burning fossil fuel [...]
B: Oh yeah it's the last few weeks.	

Figure 1: Examples of WMNs from the NeWMe corpus. WMNs that focus on *situated meaning* (left) are more typically about adjectives, while WMNs that focus on *meaning potential* (right) are more typically about nouns. For more, see the NeWMe corpus browser. E.g., *bell*: https://dev.clasp.gu.se/newme/wmn/J9P/J9P_760.

"a person is less likely to succeed if..." and the indicator is "What do you mean by success?", the trigger will be "succeed".

Here, we provide simple descriptive statistics reflecting the nature of WMN trigger phrases. This is to our knowledge the first time such data has been presented.

4.1 Concreteness of Trigger Expressions

To map mean ratings to categories that can be counted, we adopted the following categorization scheme:

- Concrete: Mean rating ≥ 3.5
- Abstract: Mean rating ≤ 2.5
- Mixed: $3.5 >$ Mean rating > 2.5
- Unknown: Word or phrase not found in the dataset

While these thresholds are to some extent arbitrary, these categories can be used to compare how concreteness relates to other categorisations as long as the same thresholds are used. The mixed effects model in Section 5 uses the raw mean ratings and is thus not affected by this choice of thresholds.

Given these category thresholds, the quantities shown in Table 4 were observed:

	NeWMe		Brysbaert	
Type	#	%	#	%
Abstract	130	33%	15,447	39%
Mixed	116	30%	10,913	27%
Concrete	120	31%	13,594	34%
Unknown	26	7%	0	0%
Total	**392**	**100%**	**39,954**	**100%**

Table 4: Distribution of concreteness

The current thresholds yield roughly similar numbers of instances per category. Applying the same thresholds to the Brysbaert et al. (2014) data yields a similar distribution to that found in the NeWMe data. This may be taken to indicate that overall (not taking into account the type of WMN), the frequency of WMNs are independent of the abstractness of the trigger phrase.

4.2 Sentiment of Trigger Expressions

We used the positivity (**PosSenti**) and negativity (**NegSenti**) metrics from SentiWordNet and classified them into Positive (**PosSenti** > 0, negativity $= 0$), Negative (**NegSenti** > 0, **PosSenti** $= 0$), Both (**PosSenti** > 0, **NegSenti** > 0) and Neither / Not Included (**PosSenti** $= 0$, **NegSenti** $= 0$ or word not included in SentiWordNet). The results are shown in Table 5. A majority of WMN triggers are neither positive nor negative, but about 1/3 are positive, negative or both (in roughly equivalent proportions).

Sentiment	#	%
Positive	50	13%
Negative	39	10%
Both	35	9%
Neither / Not Included	268	68%
Total	**392**	**100%**

Table 5: Distribution of sentiment

4.3 Other variables

The proportion of single-word vs. compound phrases (or acronyms) is seen in Table 6. About 2/3 of trigger expressions are single words, and about 1/3 are compounds[6].

Lexical Form	#	%
Single	266	68%
Compound	121	31%
Acronym	5	1%
Total	**392**	**100%**

Table 6: Distribution of lexical forms in the sample

Next, we have a look at part of speech of the WMN trigger expression head word in Table 7 where it can be noted that nouns account for almost 3/4 of trigger expressions, with adjectives and verbs at around 1/8 each.

POS	#	%
Noun	281	72%
Adjective	55	14%
Verb	50	13%
Adverb	6	2%
Total	**392**	**100%**

Table 7: Distribution of parts of speech in the sample

5 Statistical modeling

To investigate the research questions discussed in Section 1 we employ three mixed effects models.

[6]Unfortunately, we were not able to ascertain the proportion of single word vs. compound in English lexicalised expressions in the corpora used or in English in general, so it is difficult to say if trigger expressions are atypical with respect to this parameter.

The first model addresses **RQ1** by testing which interaction and trigger expression features influence the WMN type (NON or DIN). The next two models address **RQ2** by using the same features to predict which aspects of meaning (situated meaning and/or meaning potential) are negotiated in the WMN.

In all three models, we leave out items with low-frequency values for categorical variables. In particular, we filter out items whose trigger expression PoS is *adverb* and items whose WMN type is categorised as *other* (as opposed to NON or DIN). This leaves a total of 337 observations on which to base model estimates.

All three models are generalized linear mixed effects models fit by maximum likelihood estimation. The following predictor variables are used:

- `type` – whether the WMN is a NON or a DIN (not used as a predictor in the first model)
- $\texttt{conc}_\mu$ – the mean concreteness score for the trigger (lexical item or head word lemma)
- $\texttt{conc}_\sigma$ – the standard deviation (I included this because I thought it could be predictive of NON/DIN since it is essentially a measure of annotator disagreement)
- $\texttt{sent}^+/\texttt{sent}^-$ – the positive/negative sentiment scores from SentiWordNet3.0
- $\texttt{sent}^+ * \texttt{sent}^-$ – An interaction term for the positive and negative sentiment scores
- `pos` – the part of speech, coded as a one-hot (dummy) variable with noun as the reference category
- `lexform` – the compound status of the expression (single or multi-word), with `single-word` as the reference category

5.1 WMN type (NON vs. DIN)

To investigate how the variables of interest impact whether a WMN is a NON or a DIN, we use the following generalized linear mixed effects model:

$$\begin{aligned}\texttt{type} \sim 1 + \texttt{conc}_\mu + \texttt{conc}_\sigma \\ + \texttt{sent}^+ + \texttt{sent}^- + (\texttt{sent}^+ * \texttt{sent}^-) \\ + \texttt{pos} + \texttt{lexform} + (1|\texttt{corpus})\end{aligned}$$

where type is a Bernouli response variable coded with 1 for DIN and 0 for NON.

We include the source corpus (corpus) from which the WMN was drawn as a random effect variable since there are likely to be baseline differences in the propensity for NONs vs. DINs across the three corpora. The fixed effect predictors are defined as follows:

The interaction term $\texttt{sent}^+ * \texttt{sent}^-$ was included because of the way the SentiWordNet3.0 metrics are defined. The combination of the positive and negative sentiment scores can be understood as a measure of how "interested" or "sentiment-laden" the term is.

The model found $\texttt{conc}_\mu$ to have a significant negative relationship with the response variable ($\beta = -0.597$; $p = 0.0012$). We also find that *verbs* are significantly less likely to appear as the trigger expression to DINs compared to *nouns* ($\beta = -1.30643$; $p = 0.0030$). Similarly, *multi-word* trigger expressions are less likely to appear in DINs compared to *single-word* trigger expressions ($\beta = -0.785$; $p = 0.0226$).

Complete details of the models and their fit (for this and the following two models) can be found in Appendix A.

These results show that abstract trigger words are more often triggers of WMNs motivated by disagreement (DIN) than non-understanding (NON), whereas concrete triggers are more associated with non-understanding. Noun triggers are more likely to be involved in disagreements in comparison to WMNs originating in non-understanding, which are more associated with verbs. When multi-word expressions trigger a WMN, it is more likely that the WMN is a NON.

5.2 Meaning aspect

As discussed, in Section 2, WMNs can focus on the *situated meaning* of a word — what it means in *that* particular context of use — or its *meaning potential* — what it *could* mean more generally. In contrast to NON/DIN, these are not mutually exclusive (a WMN can include discussion of both types of meaning). For that reason, we model situated and potential as two separate response variables.

We don't have a specific hypothesis, so this analysis should be considered exploratory, but we decided to use the same predictors as in Section 5.1, with the addition of WMN type as a dummy-coded categorical variable, since we reason that there may be different reasons to discuss potential vs. situated meaning when there is a disagreement versus non-understanding.

The model for situated meaning is as follows:

$$\begin{aligned}\texttt{situated} \sim 1 + \texttt{type} + \texttt{conc}_\mu + \texttt{conc}_\sigma \\ + \texttt{sent}^+ + \texttt{sent}^- + (\texttt{sent}^+ * \texttt{sent}^-) \\ + \texttt{pos} + \texttt{lexform} + (1|\texttt{corpus})\end{aligned}$$

where situated is a Bernouli response variable coded with 1 if the WMN addressed the target expression's situated meaning and 0 otherwise.

We find statistically significant results for pos, with *adjectives* and *verbs* both more likely to trigger WMNs involving situated meaning than *nouns* ($\beta = 1.305$; $p = 0.0111$ and $\beta = 2.121; p = 0.0050$, respectively). The results for type are statistically non-significant, but trending negative for DINs ($\beta = -0.688; p = 0.0684$).

The model for meaning potential is analogous:

$$\begin{aligned}\texttt{potential} \sim 1 + \texttt{type} + \texttt{conc}_\mu + \texttt{conc}_\sigma \\ + \texttt{sent}^+ + \texttt{sent}^- + (\texttt{sent}^+ * \texttt{sent}^-) \\ + \texttt{pos} + \texttt{lexform} + (1|\texttt{corpus})\end{aligned}$$

We find statistically significant results for pos, with *adjectives* and *verbs* both less likely to trigger WMNs involving meaning potential than *nouns* ($\beta = -1.562$; $p = 0.0016$ and $\beta = -1.386; p = 0.0020$, respectively). The type predictor shows a statistically significant positive relationship between DIN and the focus on meaning potential ($\beta = 2.133; p < 1\mathrm{e}{-9}$).

These results show that DINs are significantly more likely than NONs to include discussion of meaning potential. The results for situated meaning are less clear, but there is some suggestion that discussions of situated meaning are more associated with NONs. In comparison to nouns, adjectives and verbs are more likely appear in discussions of situated meaning, and less likely to appear in discussions of meaning potential. Similarly, multi-word expressions are less likely to appear in discussions of meaning potential, though no clear relationship exists with situated meaning.

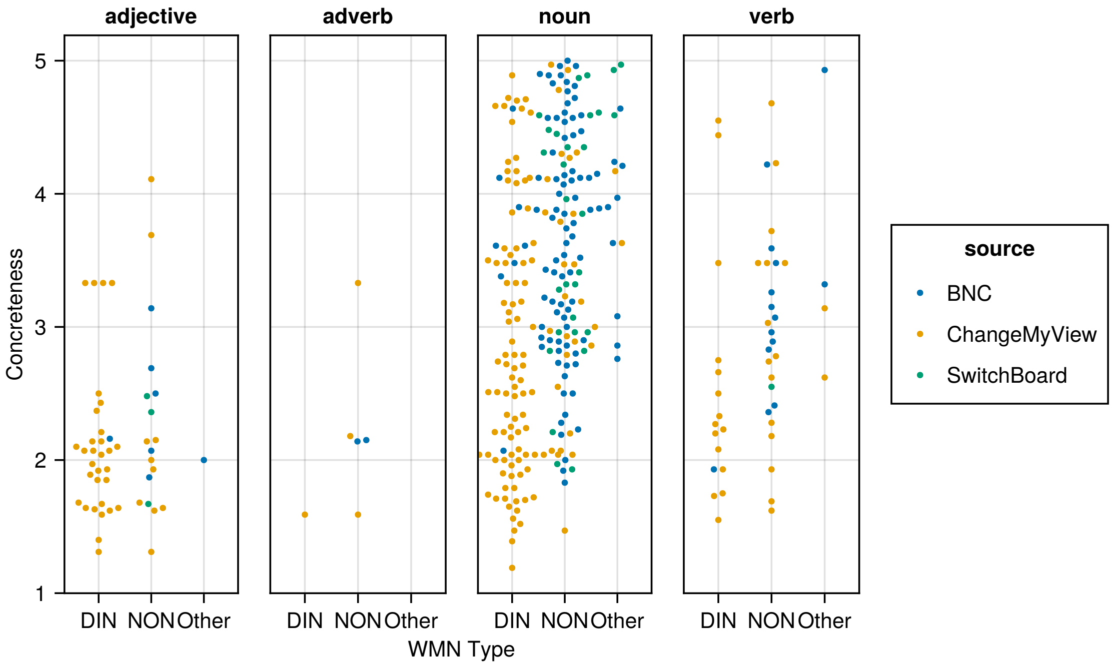

Figure 2: Violin plots of concreteness score by WMN type, broken down by part of speech.

6 Summary, discussion and future work

We reported on descriptive statistics and mixed effects modeling to understand which factors, including factors related to concreteness and sentiment in the trigger expression, affect whether the negotiation of the meaning of the term originates in non-understanding (NONs) or in disagreement (DINs), and whether the negotiation concerns situated meaning or a meaning potential.

It was found that although in general the degree of abstractness of the trigger phrase is not associated with the occurrence of WMNs overall (preliminary result), abstract trigger words are significantly more likely to result in WMN motivated by disagreement (DIN) than non-understanding (NON), whereas the converse is true of concrete trigger expressions. It could be that abstract words leave more room for individual variation in interpretation and/or that the meaning of abstract words has more implications for the idealogical goals of speakers.

A similar explanation can could be made for the clear relationship between DINs and discussions of meaning potential: Since DINs can tie in to the long-term conceptual or idealogical goals of speakers, it is more relevant to discuss what words *can* be used to mean (i.e., their meaning potential), rather than only what they mean in a particular context.

Regarding the part of speech of the trigger expressions, nouns are significantly more likely to result in DINs than verbs, and thus conversely, nouns are significantly more likely to result in NONs than verbs. Discussions of situated meaning are more likely to be triggered by adjectives and verbs than nouns. One possible explanation for this is that adjectives and verbs have more flexible meaning than nouns, allowing for more situation-specific adaptation (and potential for misalignment between speakers). Consider Figure 1 again. The WMNs of *bell* and *recent* both focus on situated meaning. In `J9P/JP_760`, speaker A uses *bell* to evoke a particular kind of situation where a bell attached to a gate or door so that it rings when someone comes in. Evidenced by this example, the situated meaning of nouns certainly *can* become the subject of WMN. However, if we compare this example with `FME/FME_18`, it's clear that the situated meaning of *recent* is unavoidably context-dependent in a way that doesn't hold for *bell*. It could be that there are systematic differences in the relationship between meaning potential and situated meaning across different parts of speech, and that this explains some of the effects we observed in Section 5.2. The mechanisms of these relationships are potential avenues for future work.

Somewhat interestingly, sentiment was not found to have any significant effect on either negotiation type or meaning aspect. Note that sentiment may still be a predictor for the occurrence of WMNs overall; this has not been investigated here.

In future work, we would like to investigate factors which influence whether a word is likely to be the topic of a WMN. This would require data describing the NeWMe source corpora and/or English in general along the dimensions we have used here. Here, we could only do this in a preliminary manner (in Section 3.2) thanks to the existence of existing data about abstractness in English lexical items.

As always, more data would provide a better basis for analysis. Classifying data according to dialogue genre Ginzburg (2012) and/or activity type Allwood (1987) would help tease these factors apart from the Interaction Type which currently conflates them with the medium of communication (spoken or written interactions). Another obvious extension is to see if the results reported here are the same in other languages.

Acknowledgments

This work was supported by the Swedish Research Council (VR) grant 2022-02125 *Not Just Semantics: Word Meaning Negotiation in Social Media and Spoken Interaction*, and VR grant 2014-39 for the establishment of the Centre for Linguistic Theory and Studies in Probability (CLASP) at the University of Gothenburg.

References

Jens Allwood. 1987. On the analysis of communicative action. In M. Brenner, editor, *The Structure of Action*. London: Basil Blackwell.

Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. 2010. SentiWordNet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta. European Language Resources Association (ELRA).

Marc Brysbaert, Amy Beth Warriner, and Victor Kuperman. 2014. Concreteness ratings for 40 thousand generally known english word lemmas. *Behavior research methods*, 46:904–911.

Aina Garí Soler, Matthieu Labeau, and Chloé Clavel. 2023. Measuring lexico-semantic alignment in debates with contextualized word representations. In *Proceedings of the First Workshop on Social Influence in Conversations (SICon 2023)*, pages 50–63. Association for Computational Linguistics.

Aina Garí Soler, Jenny Myrendal, Chloé Clavel, and Staffan Larsson. 2025. The newme corpus: A gold standard corpus for the study of word meaning negotiation.

Aina Gari Soler, Jenny Myrendal, Chloé Clavel, and Staffan Larsson. forthcoming. The newme corpus: A gold standard corpus for the study of word meaning negotiation. *Submitted for review*.

Jonathan Ginzburg. 2012. *The Interactive Stance*. Oxford University Press, New York.

Staffan Larsson and Jenny Myrendal. 2024. Not just semantics: Word meaning negotiation in social media and spoken interaction. In *Proceedings of the 2024 CLASP Conference on Multimodality and Interaction in Language Learning*, pages 56–61.

Jenny Myrendal. 2015. *Word Meaning in Interaction: Semantic Negotiation in Online Forums*. Phd thesis, University of Gothenburg.

Jenny Myrendal. 2019. Negotiating meanings online: Disagreements about word meaning in discussion forum communication. *Discourse Studies*, 21(3):317–339.

Jenny Myrendal. 2025. Repair of claimed non-understanding of word meaning in online discussion forum interaction. *Dialogue & Discourse*, 16(1).

Jenny Myrendal and Staffan Larsson. 2025. Semantic conflict in online discussions: Negotiating the meaning of 'lying'. *Journal of Language Aggression and Conflict*.

Kerstin Noren and Per Linell. 2005. Meaning potentials and their empirical substantiations. In *Paper presented at the 9th International Pragmatics Conference*.

Kerstin Norén and Per Linell. 2007. Meaning potentials and the interaction between lexis and contexts: An empirical substantiation. *Pragmatics. Quarterly Publication of the International Pragmatics Association (IPrA)*, 17(3):387–416.

James Pustejovsky and Nikhil Krishnaswamy. 2020. Situated Meaning in Multimodal Dialogue: Human-Robot and Human-Computer Interactions. 61(3):17–41.

Evangeline Marlos Varonis and Susan Gass. 1985. Non-native/non-native conversations: A model for negotiation of meaning. *Applied linguistics*, 6(1):71–90.

A Statistical model results

Details for `type` model

$$\texttt{type} \sim 1 + \texttt{conc}_\mu + \texttt{conc}_\sigma + \texttt{sent}^+ + \texttt{sent}^- + (\texttt{sent}^+ * \texttt{sent}^-) + \texttt{pos} + \texttt{lexform} + (1|\texttt{corpus})$$

The model was fit by maximum likelihood estimation (nAGQ$=9$) with a Bernouli response variable and logistic linking function. The fit is as follows:

logLik	deviance	AIC	AICc	BIC
-139.1763	278.2902	298.3526	299.0401	336.3738

The random effect for WMN `source` had variance 7.17901 and standard deviation 2.67937. The details for the fixed effects were as follows:

	Coef.	**Std. Error**	**z**	**Pr(>|z|)**
(Intercept)	0.353722	1.93684	0.18	0.8551
$\texttt{conc}_\mu$	-0.596865	0.184348	-3.24	0.0012
$\texttt{conc}_\sigma$	-0.204307	0.564173	-0.36	0.7173
$\texttt{sent}^+$	0.664035	1.15001	0.58	0.5637
$\texttt{sent}^-$	0.199255	1.1104	0.18	0.8576
`pos`: *adjective*	-0.727312	0.507808	-1.43	0.1521
`pos`: *verb*	-1.306430	0.439714	-2.97	0.0030
`lexform`: *multi-word*	-0.784656	0.344007	-2.28	0.0226
$\texttt{sent}^+ * \texttt{sent}^-$	-5.356570	5.5816	-0.96	0.3372

Details for `situated` model

$$\texttt{situated} \sim 1 + \texttt{type} + \texttt{conc}_\mu + \texttt{conc}_\sigma + \texttt{sent}^+ + \texttt{sent}^- + (\texttt{sent}^+ * \texttt{sent}^-) + \texttt{pos} + \texttt{lexform} + (1|\texttt{corpus})$$

The model was fit by maximum likelihood estimation (nAGQ$=9$) with a Bernouli response variable and logistic linking function. The fit is as follows:

logLik	deviance	AIC	AICc	BIC
-161.2996	322.5839	344.5992	345.4268	386.4225

The random effect for WMN `source` had variance 0.328711 and standard deviation 0.573333. The details for the fixed effects were as follows:

	Coef.	Std. Error	z	Pr(>|z|)
(Intercept)	1.58082	1.02038	1.55	0.1213
type: DIN	-0.687528	0.377333	-1.82	0.0684
$\texttt{conc}_\mu$	-0.00794446	0.16794	-0.05	0.9623
$\texttt{conc}_\sigma$	-0.332505	0.496781	-0.67	0.5033
$\texttt{sent}^+$	-1.52334	1.00917	-1.51	0.1312
$\texttt{sent}^-$	-0.345484	0.936351	-0.37	0.7122
pos: *adjective*	1.30509	0.513597	2.54	0.0111
pos: *verb*	2.12148	0.755799	2.81	0.0050
lexform: *multi-word*	0.360319	0.322338	1.12	0.2636
$\texttt{sent}^+ * \texttt{sent}^-$	5.32303	5.34836	1.00	0.3196

Details for potential model

$$\texttt{potential} \sim 1 + \texttt{type} + \texttt{conc}_\mu + \texttt{conc}_\sigma + \texttt{sent}^+ + \texttt{sent}^- + (\texttt{sent}^+ * \texttt{sent}^-) + \texttt{pos} + \texttt{lexform} + (1|\texttt{corpus})$$

The model was fit by maximum likelihood estimation (nAGQ= 9) with a Bernouli response variable and logistic linking function. The fit is as follows:

logLik	deviance	AIC	AICc	BIC
-161.2996	322.5839	344.5992	345.4268	386.4225

The random effect for WMN source had variance 0.240946 and standard deviation 0.490862 The details for the fixed effects were as follows:

	Coef.	Std. Error	z	Pr(>|z|)
(Intercept)	-0.0314509	1.04118	-0.03	0.9759
type: DIN	2.13327	0.342115	6.24	$< 1e{-}9$
$\texttt{conc}_\mu$	-0.24192	0.172902	-1.40	0.1618
$\texttt{conc}_\sigma$	0.217801	0.497279	0.44	0.6614
$\texttt{sent}^+$	1.62563	0.990182	1.64	0.1006
$\texttt{sent}^-$	0.0794144	0.975404	0.08	0.9351
pos: *adjective*	-1.56217	0.49569	-3.15	0.0016
pos: *verb*	-1.38621	0.447837	-3.10	0.0020
lexform: *multi-word*	-0.550717	0.304911	-1.81	0.0709
$\texttt{sent}^+ * \texttt{sent}^-$	1.48935	5.80966	0.26	0.7977