Title
The categorization of intransitive verbs in Mandarin: Evidence from word2vec modeling and behavioral experiment

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Authors
Lin, Kaiying
Washington, Peter

Publication Date
2023

Peer reviewed

The categorization of intransitive verbs: Evidence from word2vec modeling and a behavioral experiment

Kaiying Lin (kylin@hawaii.edu)
Department of Linguistics & Information and Computer Science, University of Hawai'i, Mānoa

Peter Y. Washington (pyw@hawaii.edu)
Department of Information and Computer Science, University of Hawai'i, Mānoa

Abstract

Intransitive verbs in human language can be subcategorized into at least two linguistic categories, unaccusative and unergative verbs. The categorization of these verbs has been a subject of debate, with Projectional approaches emphasizing the role of verb semantics and Constructional approaches highlighting the importance of the sentence context. We utilize a word2vec model to capture the environmental influence on verb semantics, providing evidence that supports both approaches. Our results demonstrate that the categorization of Unaccusativity for new verbs can be influenced by the sentence context in which they appear, and the frequency of verbs in specific contexts also plays a significant role. Additionally, through a child language acquisition experiment, we show that the sentence environment has a significant impact on the categorization of Unaccusativity when the semantics of new verbs are provided. These findings suggest that both approaches have merit, highlighting the psychological universality of Unaccusativity across languages.

Keywords: intransitive verbs, Unaccusativity, Projectional approach, Constructional approach, word2vec, computational modeling

Introduction

Linguists have long debated the existence of two linguistic categories of intransitive verbs in language, namely, unergative and unaccusative verbs. When the subject of an intransitive verb shows grammatical and semantic properties similar to the subjects of transitive verbs, it is an unergative verb; when it shows similar properties to the objects of transitive verbs, it is an unaccusative verb. These categories are determined using specific diagnostics. For instance, in English, the grammaticality of -er is a diagnostic to divide intransitive verbs. If the subject of an intransitive verb can take the -er suffix, it is considered unergative verb; otherwise it is classified as unaccusative (e.g., (1)).

(1)Unergative verbs: runner, singer, swimmer, listener
Transitive verbs: eater, drinker, player, driver
Unaccusative verb: *dier, *comer, *disappearer, *arriver

This observation is formalized as the Unaccusativity hypothesis (Permultter, 1978), and two approaches were proposed to investigate the underlying factors that determine categorization in language. Projectional approaches (Chomsky, 1981; Hale & Keyser, 1993; Levin & Rappaport Havov, 1995; Pinker, 1989) emphasize that the meaning of the verb plays a more definitive role in deciding the category of the verb, while Constructional approaches (Hoekstra, 1992; Borer, 1994; Goldberg, 1995; van Hout, 1996; Ritter & Rosen, 1996) stress the sentential environment in which the verb appears decides the category of the verb. The Projectional approaches received more attention in linguistics due to the relative simplicity of manipulating semantic factors in theoretical and experimental studies. In (1), one can propose that unergative verbs involve atelic and agentative events, but unaccusative events involve telic events and a meaning of change. On the other hand, to determine the final category of a verb, Constructional approaches require an examination of its occurrences in natural language and an identification of the specific environments that give rise to unaccusative or unergative verbs.

Recent advances in neural network models based on distributional semantics have contributed to the development

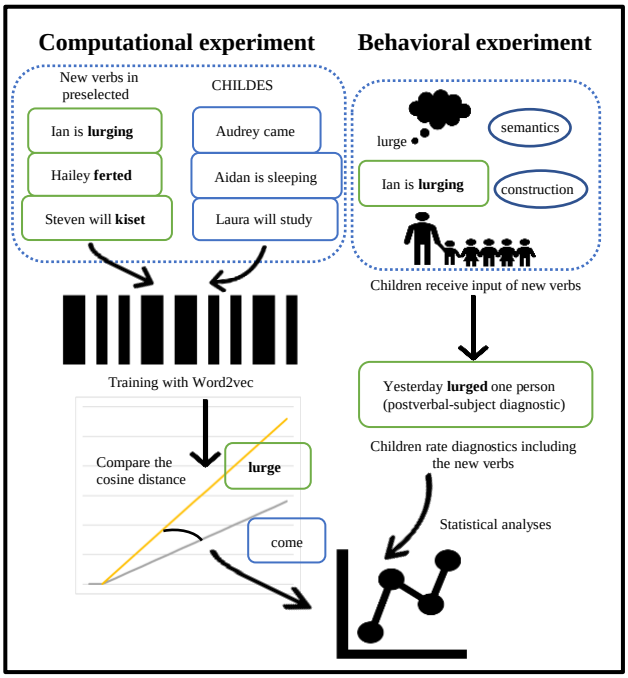

Figure 1: The experimental procedures for both the computational and behavioral experiments. The computational experiment involves generating and comparing numerical representations of verbs. The behavioral experiment involves the child participants rating diagnostics with new verbs after receiving different kinds of input.

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society.*

of Constructional approaches by leveraging the associations between a word and its context within a sentence. These models provide insights into the relationship between a verb and its surrounding words, aiding in the understanding of Constructional approaches.

One of many possible embedding models, word2vec (Mikolov et al., 2013) is a shallow neural network that has proven efficacy across of variety of natural language processing tasks. The model is fed with sentences in the corpus and each word in the corpus is encoded as a vector. During training, the model adjusts the weights between layers based on the associations between the target word and its surrounding words, gradually learning the distribution of the input corpus. The resulting word embeddings represent words as vectors, with similar words being located closer to each other in the vector space. This proximity reflects their shared semantic and contextual characteristics in natural language. Hence, the distance in vector space can capture the semantic (dis)similarities in natural language. This makes the word2vec model a useful tool to investigate how the environment can form the semantics of a verb and hence can affect categorization.

The efficacy of the word2vec model can facilitate language-related research since the input to a model can be manipulated without longitudinal engagement from participants. This enables explicit observation of lexical variation after receiving priming (e.g., Ettinger & Linzen, 2016) and stimulus (e.g., Ettinger et al., 2016). In the domain of language acquisition, word2vec can remove the burden of searching for participants in experimental protocols. It serves as a practical tool to manipulate input and study the resulting relationships between word meanings in the input and output (e.g., Weber & Colunga, 2022; Fourtassi, 2020). Moreover, the efficient nature of the computational modeling provides a preliminary result that can be further examined in follow-up behavioral experiments with real people.

In this study, we examine whether the categorization of unaccusative vs. unergative can stem from the semantics or the environment of the verb. We develop a word2vec model to mimic the environment of Mandarin-speaking children and to examine how the environment of the sentence can influence semantics so as to affect the categorization of Unaccusativity of intransitive verbs. The results suggest that new verbs (verbs that have zero occurrences except in pre-selected sentences) in some pre-selected sentences show greater cosine distances to existing verbs in one category and hence were seen as belonging to that category. A further manipulation of the occurrence of these pre-selected sentences demonstrates that the frequency of occurrence can also be influential in pulling verbs into one category. Our findings provide evidence that the environments and frequency of verbs can affect the semantic embeddings of verbs and ultimately influencing the categorization.

We further confirm the above result in a child language acquisition behavioral experiment. Two groups of children are tested regarding how they decided the category of Unaccusativity of new verbs with or without pre-selected sentences when both groups are provided the semantics of new verbs. We find that the group provided with pre-selected sentences rate the diagnostic sentences significantly closer to the expected values. This provides additional support for the hypothesis that the environment of verbs is indeed a factor in the categorization.

The two experiments in this study suggest that both Projectional and Constructional approaches require consideration. The modeling of word2vec reveals that the environment can affect the semantic interpretation of verbs and the result of categorization. The behavioral experiment further suggests that the environment can reinforce the categories of Unaccusativity of verbs in addition to verbal semantics. These results imply that the incorporation of both Constructional and Projectional approaches should be considered for a unified theory of Unaccusativity (e.g., Rappaport Havov & Levin, 1998; Sorace, 2000).

Computational Experiment

Methods

We used pre-selected sentences to represent the grammatical and semantic environment that we hypothesized can affect the Unaccusativity of the verbs within them. To investigate the impact of these pre-selected sentences on the categorization process, we trained a word2vec model using the CHILDES corpus, which contained a collection of child-directed speech along with the pre-selected sentences. We then compared the cosine distance similarity between the 22 new verbs in the pre-selected sentences and the existing verbs found in the corpus. The existing verbs were chosen based on their presence in postverbal subject or durative sentences, widely recognized as diagnostics for Unaccusativity in Mandarin (Huang, 1987; Liu, 2007). We used the comparison with existing verbs to observe whether (dis)similarities will form two or more clusters, as presented in Table 1.

Pre-selected sentences

Table 1 presents the 11 types of pre-selected sentences used in this study: LVS, resultative (2nd), *ba*-sentence, perf (null-subject), perf, resultative (1st), resultative (1st-null), will+V, V+dur, want+V, and imperf+V. These sentences served as contexts in which intransitive verbs can occur, and we classified each verb into one of three categories: unergative, unaccusative, or neutral, based on their likelihood of influencing the verbs within those sentences. This classification was informed by previous literature, considering factors such as the position of the new verbs, agentivity, and telicity (Li & Thompson, 1981; Tsang, 1981; Huang, 2006; Liu, 2007; Wang, 2010; Laws & Yuan, 2010). For each pre-selected sentence type, we created two sentences with two new verbs. Each sentence was increased to 4 occurrences with separate verbs to yield 88 pre-selected sentences in total, producing a larger sample pool. To ensure the reliability of results, we employed bootstrapped sampling, randomly selecting nouns from three lists containing 30 different subjects, 10 different objects, and 10 different

locations. See Table 1 for the expected categories of Unaccusativity and Table 2 for examples and translations of pre-selected sentences.

Furthermore, the occurrence of each sentence was increased to 4, 8, 32 times with nouns generated via bootstrapped sampling. The three types of occurrences of pre-selected sentences were fit to three different models (low, middle, and high) along with CHILDES corpora.

Table 1: Pre-selected sentences and their expected categories.

Sentence No.	Pre-selected sentence	Expected effects of categories
1	LVS	unaccusative
2	Resultative (2nd)	unaccusative
3	Ba-sentence	unaccusative
4	Perf (null-subject)	unaccusative/neutral
5	Perf	unaccusative/neutral
6	Resultative (1st)	neutral
7	Resultative(1st-null)	neutral
8	Will+V	neutral/unergative
9	V+Dur	unergative
10	Want_V	unergative
11	Imperf+V	unergative

Experiment

We trained three word2vec CBOW models using the Gensim package (Rehurek & Sojka, 2011) in Python. We used the Taiwanese Mandarin corpora in CHILDES, i.e., Chang1, Chang2, ChangPlay, ChangPN, TCCM, TCCM-Readings. A total of 173,673 sentences from the corpora plus the pre-selected sentences were used for training. We split sentences into words using the jieba package (Jieba, 2022) and tokenized the words using a self-defined dictionary. To avoid randomization, we hardcoded the parameters of the word2vec training process. To ensure reliable results, we trained each model (low, middle, high) with different occurrences of pre-selected sentences a total of 20 times. We extracted the mean word vectors of 22 new verbs and compared with two existing unaccusative verbs with respect to their cosine distance, i.e., *diao4* 'drop', *lai2* 'come', which occurred in the postverbal-diagnostic sentence in the CHILDES corpora. The word vectors were next compared to existing unergative verbs that occurred in the durative diagnostic in the corpora, i.e., 'jump' *tiao4*, 'play' *wan2*, 'cry' *ku1*. The above comparisons imply that children at least are aware of the Unaccusativity of those existing verbs in the corpora, and can use those existing verbs to define the categories of new verbs by comparing their semantics. The success of the experiment relied on the assumption that the environment was a crucial factor, and the different occurrences between unergative and unaccusative pre-selected sentences gave rise to an obvious categorization. Additionally, it also depended on whether the (dis)similarities measured by cosine distance in the vector space could effectively reflect the distinction between unaccusative and unergative verbs.

Table 2: Pre-selected sentences, glossing, and translation. (Subj, obj, loc are bootstrapped sampled and the positions of new verbs are in **bold**).

No.	Pre-selected sentences	Glossing of Mandarin sentences[1]	Example and Translation in English
1	LVS	loc **V**-perf subj	'Down (loc) **came** the rain (subj)'
2	Resultative (2nd)	subj V1 obj **V**-perf	'Teachers (subj) beat a person (obj) (until he gets) **hurt**'
3	Ba-sentence	subj ba obj **V**-perf	'Teachers (subj) beat a person (obj) (until he gets) **hurt**'
4	Perf(null-subject)	(subj)**V**-perf	'(Father) **left!**'
5	Perf	subj **V**-perf	'The father (subj) **left**'
6	Resultative (1st)	subj **V** V2-perf	'Gasoline (subj) **ran** out'
7	Resultative (1st-null)	(subj) **V** V2-perf	'(Birds) **ran** away'
8	Will+V	subj will **V**	'Animals (subj) always will **go**'
9	V+Dur	subj **V**-dur	'People (subj) are **running**'
10	Want_V	subj want to **V**	'Boys (subj) want to **go**'
11	Imperf+V	subj keep **V**	'Father (subj) always keeps **running**'

Results

Figures 2 and 3 depict the mean and standard deviation of the cosine similarities of new verbs. The comparison with existing unaccusative verbs is depicted in Figure 2 and that with existing unergative verbs in Figure 3. Based on the expected categories of Unaccusativity in pre-selected sentences in Table 1, we rearrange the pre-selected sentences to place unaccusative sentences on the left, unergative sentences on the right, and neutral sentences in the middle. We hypothesize that the cosine similarities (ranging from -1 to 1) will show a left-slanted slope in Figure 2 and a right-slanted slope in Figure 3. The results validate our hypothesis: new verbs in pre-selected sentences #1,2,3 show more cosine similarities in Figure 2 and less cosine similarities in Figure 3. New verbs in #9,10,11 show the contrary trend and verbs #4,5,6,7,8 show intermediate cosine similarity values. The

[1] Abbreviation: perf=perfective marker (*-le* in Mandarin), dur=durative marker (*-zhe*), imperf=imperfective marker (*zai*).

difference between the cosine similarities of new verbs in #1,2,3 and #9,10,11 are statistically significant in comparison with existing unaccusative verbs and unergative verbs (Student t-test, t(34)=7.03, p<.001***; t(34)=-7.77, p<.001***), showing a successful and clear categorization. Word vectors are regarded as the semantic representation of verbs, and this finding firstly demonstrates the environment of pre-selected sentences can affect the semantics of new verbs in them, which eventually affects the categories of Unaccusativity.

We also find that the categories of these pre-selected sentences not only meet our expected categories of Unaccusativity but also show varying categories. Even within the same expected categories, the environmental effects on new verbs are not uniform. In Figure 2, some unaccusative sentences (e.g., #1) have a higher effect in categorizing verbs than others (e.g., #2, #3). The same relative ordering is seen with unergative sentences (e.g., #11> #10). The different levels of effects from a variety of sentences suggest that the environmental effect is likely to be more complex than the expected two or three categories in Table 1, and likely to form a gradient pattern.

Furthermore, we see in both graphs that with more occurrences of the new verbs, the cosine similarities of the verbs are inclined to stabilize. By contrast, the similarities tend to fluctuate with fewer occurrences. When training data with fewer occurrences of new verbs and pre-selected sentences (low model in Figure 2, 3) are included, we observe a jagged line of cosine similarities, while the lines form smoother slopes when more occurrences of new verbs and pre-selected sentences are given (middle and high model in Figures 2 and 3). The means of the standard deviation of the three models decreases (low, middle, high, std=0.05, 0.03. 0.02) in comparison with existing unaccusative and unergative verbs. These differences in the standard deviations among models reach statistical significance (low vs middle, t(20)=2.60, p=.02*; low vs high, t(20)=5.17, p<.001***; middle vs high, t(20)=4.57, p<.001***), and a similar trend is observed in comparison with existing unergative verbs as well (low vs middle, t(20)=2.07, p=.04*; low vs high, t(20)=5.17, p=<.001***; middle vs high, t(20)=4.02, p<.001***). These findings suggest more occurrences of the pre-selected sentences can result in clearer clustering of

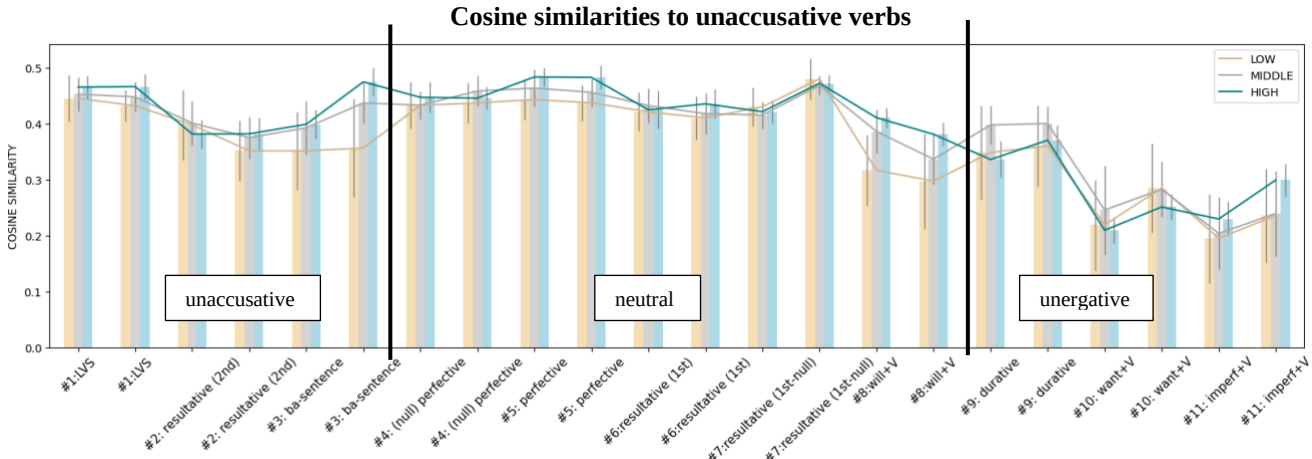

Figure 2: Average cosine similarities of new verbs compared to existing unaccusative verbs. LOW, MIDDLE and HIGH models have varying frequency of pre-selected sentences and new verbs. The Figure is divided into three partitions, each representing a sentential category as shown in the box. (Same for Figure 3).

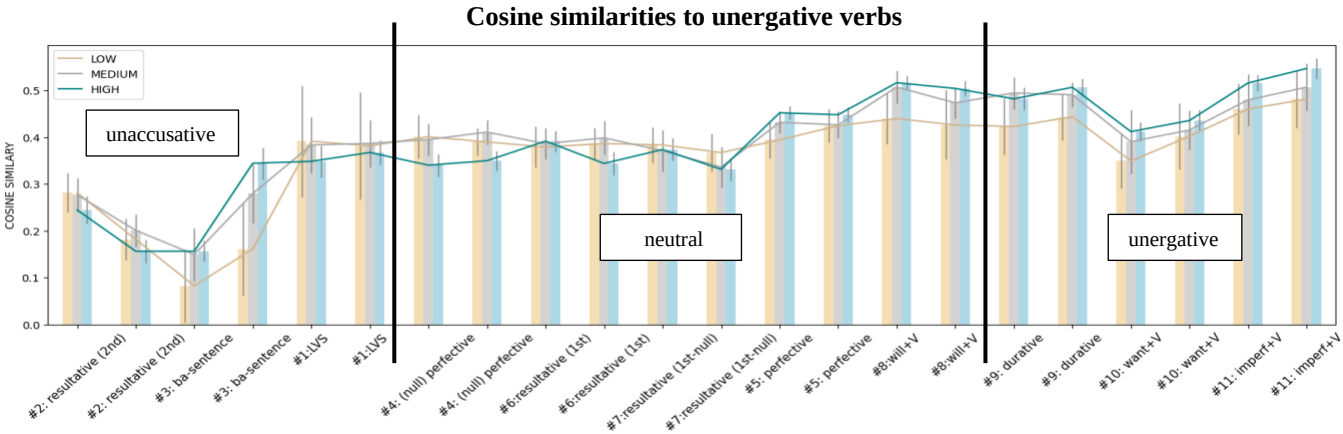

Figure 3: Average cosine similarities of new verbs compared to existing unergative verbs.

unergative and unaccusative verbs, and frequency is a factor in the categorization of Unaccusativity.

Behavioral Experiment

We perform a behavioral experiment to serve as additional support for the results of the previous computational experiment. We examine how the environment in which verbs appear can impact children's categorization of Unaccusativity for new verbs in real-world contexts.

Methods

This experiment consisted of two groups, namely the non-construction group and the construction group, and employed an acceptability judgment task based on Ambridge (2012). Unlike the computational experiment, where verbs were presented without their associated semantics, this behavioral experiment aimed to examine the environmental effect by comparing the two groups. In the non-construction group, only the semantics of the new verbs were provided, while the construction group received both verbal semantics and pre-selected sentences from the same category. The objective was to assess the extent to which the categorization of Unaccusativity could be improved. The improvement was measured by comparing the ratings of diagnostic sentences between the two groups. Our assumption is that the inclusion of pre-selected sentences in the construction group can significantly enhance the categorization of new verbs if the verb's environment influences categorization.

Procedure

Child participants and the investigator sat together on one side of the table and watched PowerPoint slides. They were informed that their task was to assist a dog in learning Mandarin by rating the acceptability of sentences. Even if they were unfamiliar with a verb, they were instructed to rely on their intuition to help the dog. In the non-construction group, participants received animations (semantics) as input. In the construction group, participants received both animations (semantics) and pre-selected sentences containing new verbs (construction). A slide with a circle (representing "acceptable") and a cross (representing "unacceptable") then appeared on the screen. Participants listened to diagnostic sentences formed with new verbs and made a binary choice to indicate the acceptability of the sentence. Before the official experiment began, four practice trials were conducted in the same setting. These trials included two items with existing verbs and two items with new verbs. Participants had to answer all the practice trials correctly before proceeding to the official experiment. If any mistakes were made, the practice trials were repeated. The entire experiment lasted approximately 15-20 minutes.

Participants

Sixty-six Mandarin-speaking children aged 4;10 to 6;4 (mean=5;6) were recruited from a kindergarten located in Taipei, Taiwan.The participants were evenly divided into two groups. Children aged around 6yrs were selected because Lin & Deen (2021) suggests children do not exhibit the same level of differentiation ability in Unaccusativity as adults until approximately 6 years of age.

Materials

The experiment utilized six new verbs, comprising two unergative-targeted verbs and four unaccusative-targeted verbs in both the non-construction and construction groups. Due to length constraints, only three pre-selected sentences (#1, #2 for unaccusative, and #11 for unergative) were included in the construction group. These sentences were identified as highly effective in categorization based on the previous computational experiment. Consequently, the pre-selected sentence #11 contained two unergative-targeted verbs, while two out of the four unaccusative-targeted verbs were put in a respective pre-selected sentence (#1 or #2), and the other two unaccusative-targeted verbs were present in both pre-selected sentences (#1, #2). The inclusion of items with two pre-selected sentences aimed to assess the potential additive constructional effect when two such sentences were provided. To evaluate categorization outcomes, these six new verbs were crossed by two diagnostics (postverbal-subject and durative diagnostics), resulting in 12 critical items. Five filler items consisting of existing verbs were inserted after every two critical items.

The semantics of the verbs, represented through animations, were created based on telicity, which is the most significant semantic distinction between unaccusative and unergative verbs. Telicity indicates whether a verb inherently has an endpoint in the event it expresses and is considered crucial in Mandarin Unaccusativity, as argued by Liu (2007) and Lu (2019). To prevent participants from substituting the new verbs with existing ones, the animations for the new verbs were designed to include at least two transitive actions, while still conveying the semantics of the two semantic subclasses (telic or atelic). For example, the animation for the new unaccusative verb *zhi4* depicted actions such as "hit by thunder" and "become small," presenting a telic scenario where "a girl is hit by the thunder and then becomes small." The selection of the six new verbs was based on a norming test involving five native Taiwanese Mandarin-speaking adults. These adults watched the animations associated with the new verbs and selected existing verbs that best matched the animations. Three existing verbs were provided as options for each animation, and none of the existing verbs associated with the selected six animations were chosen more than three times. This result suggests that most adults could not directly associate the animations of these new verbs with any existing verbs, demonstrating the distinguishability of the new verbs from existing ones.

Prediction

If the pre-selected sentences are effective at reinforcing the categorization of new verbs, we would expect to see significantly closer ratings to the expectations on diagnostic sentences in the construction group compared to the non-

construction group. Specifically, the durative diagnostic should elicit more rejection for unaccusative-targeted new verbs and more acceptance for unergative-targeted new verbs in the construction group compared to the non-construction group. Conversely, the postverbal-subject diagnostic should be acceptable for unaccusative-targeted new verbs but not for unergative-targeted new verbs. However, if there is no constructional effect, we would not anticipate any statistically significant improvement in the construction group compared to the non-construction group.

Results

We conducted logistic regression analyses using Python's statsmodels (Seabold & Perktold, 2010) to examine the data from the two groups. The models included the independent variables of unerg/uancc (the semantic category of the verb) and diagnostic sentences, with children's ratings as the dependent variable. The results revealed statistically significant interaction effects between the two independent variables in both models (z=-2.702, p<.001***; z=2.43, p<.05*), indicating that children were able to differentiate between the two types of verbs using the two diagnostics in both groups. This finding supports the idea that semantics plays a role in categorization, which aligns with Projectional approaches.

We proceeded to transform children's ratings into binary measures of 'accuracy,' labeling ratings as 'accurate' when they matched the expected ratings and as 'inaccurate' otherwise. Figure 5 presents the distribution of 'accuracy' in both the construction group and the non-construction group. Notably, the construction group exhibited higher numerical 'accuracy' compared to the non-construction group. In a logistic regression analysis (using Python's statsmodels, Seabold & Perktold, 2010), with 'accuracy' as the dependent variable and verb_type (representing different verbal categories) and group as independent variables, the group variable was a statistically significant predictor (z=3.781, p<.001***). Subsequently, we employed the same logistic regression model, with different pre-selected sentences as the independent variable and 'accuracy' as the dependent variable. As depicted in Figure 5, all the pre-selected sentences (#1, #2, #11, and the combination of #1 and #2) were significant predictors of 'accuracy' (#1: z=2.148, p=.032<.05*; #2: z=3.637, p<.001***; #1+2: z=-2.11, p=.035<.05*; #11: z=2.367, p=.018<.05*). These findings lend support to the Constructional approaches, suggesting that the context provided by pre-selected sentences can enhance the categorization of verbs.

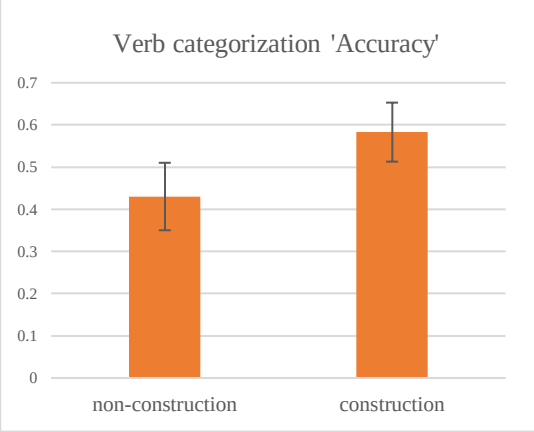

Figure 4: The 'accuracy' of the non-construction and construction groups (n=33 in each).

Conclusion

This paper contributes to the examination of two approaches to Unaccusativity within the context of language acquisition. The key contributions of our study are as follows: **(1)** We demonstrate that the environment (specifically, the sentence) in which a verb appears plays a decisive role in determining the categorization of Unaccusativity in the computational experiment. **(2)** Our findings indicate that the more frequently a verb appears in effective pre-selected sentences, the more pronounced its categorization becomes. **(3)** We observe a gradient effect of Unaccusativity from pre-selected sentences. **(4)** Through a behavioral experiment with children, we show that the sentences in which a verb appears can reinforce the categorization of Unaccusativity. **(5)** Our results reaffirm the significance of verbal semantics in Unaccusativity, supporting the arguments put forth by Projectional approaches. This is evidenced by the use of semantic representations of word vectors in the computational experiment and the observed effect of verb semantics on categorization in the behavioral experiment. In conclusion, our study not only provides support for the influence of the environment on Unaccusativity categories (Constructional approaches), but also strengthens the arguments made by Projectional approaches. This suggests that both approaches are essential in understanding the nature of Unaccusativity.

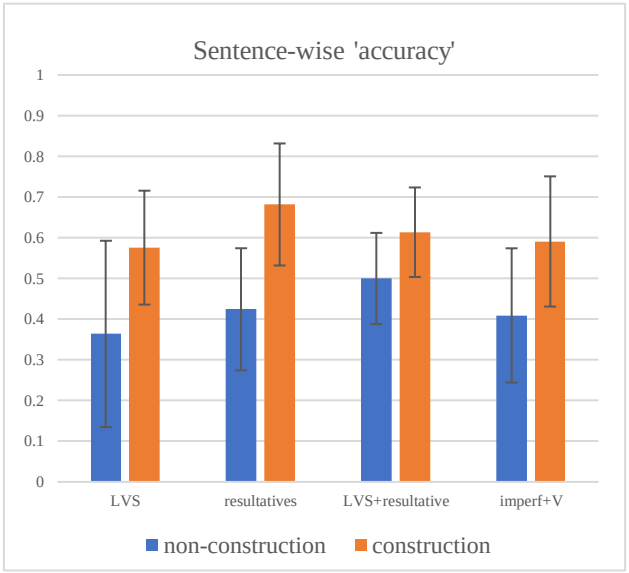

Figure 5: The 'accuracy' of different sentences in the non-construction and construction groups.

Acknowledgments

We would like to express our gratitude to Aidan Coleman, Laura Griffin, and Audrey Lai for the proofreading.

References

Ambridge, B. (2012). Assessing grammatical knowledge (with Special Reference to the Graded Grammaticality Judgment Paradigm). In E. Hoff (Ed.), *Research methods in child language: A practical guide* (pp.113-132). Hoboken, NJ: Wiley-Blackwell.

Borer, H. (1994). The projection of arguments. In E. Benedicto and J. Runner (Eds.), *University of Massachusetts Occasional Papers in Linguistics* 17, University of Massachusetts.

Chomsky, N. (1981). *Lectures on Government and Binding*. Hawthorne, NY: Foris.

Ettinger, A., Feldman, N. H., Resnik, P., & Phillips, C. (2016). Modeling N400 amplitude using vector space models of word representation. In *Proceedings of the 38th Annual Conference of the Cognitive Science Society*.

Ettinger, A., & Linzen, T. (2016). Evaluating vector space models using human semantic priming results. In *Proceedings of the 1st Workshop on Evaluating Vector-Space Representations for NLP*, 72–77. Association for Computational Linguistics.

Fourtassi, A. (2020). Word Co-occurrence in Child-Directed Speech Predicts Children's Free Word Associations. *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics*, 49–53. Association for Computational Linguistics.

Goldberg, A. E. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: Chicago University Press.

Hoekstra, T. (1992). Aspect and Theta-theory. In Roca, I. M. (Ed.) *Thematic Structure: Its Role in Grammar* (pp.145-174). Berlin: Mouton de Gruyter.

Hale, K., & Keyser, S. J. (1993). On the argument structure and the lexical expression of Syntactic Relations. In Hale, K. & S. J. Keyser. (Eds.), *The View from Building Essays in Honor of Sylvain Bromberger* (pp. 53-109). Cambridge, MA: MIT Press.

van Hout, Angeliek. (1996). *Event semantics of verb frame alternations. A case study of Dutch and its acquisition*. Tilburg: University of Tilburg.

Huang, J. C.-T. (1987). Existential sentences in Chinese and (in)definiteness. In E. J. Reuland & A. G. B. Ter Meulen (Eds.), *The Representation of (In)definiteness* (pp.226-253). Cambridge, MA: MIT Press.

Huang, C.-T. James. (2006). Resultatives and unaccusatives: a parametric view. *Bulletin of the Chinese Linguistic Society of Japan* 253, 1-43.

Laws, J., & Yuan, B. (2010). Is the core-peripheral distinction for unaccusative verbs cross-linguistically consistent? Empirical evidence from Mandarin. *Chinese Language and Discourse*, 1(2), 220–263. DOI: 10.1075/cld.1.2.03law

Levin, B., & Rappaport-Hovav, M. (1995). *Unaccusativity: At the syntax-lexical semantics interface*. Cambridge, MA: MIT Press.

Li, C., & Thompson, S. (1981). *Mandarin Chinese: A Functional Reference Grammar*. Berkeley, CA: University of California Press.

Lin, K., & Deen, K. (2021). Unaccusativity in Mandarin Child Language. In D. Dionne, & L.-A.V. Covas (Eds.), *Proceedings of the 45th Boston University Conference on Language Development*, 486-498. Somerville, MA: Cascadilla Press.

Liu, F. (2007). Auxiliary selection in Chinese. In R. Aranovich (Eds.), *Split auxiliary systems* (pp.181–205). J Amsterdam: John Benjamins.

Lu, Y. (2019). *The acquisition of Unaccusativity in Mandarin Chinese*. Doctoral Dissertation. The Chinese University of Hong Kong.

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk. Third Edition.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781. (low, middle and high indicate the three models with different occurrences of).

Perlmutter, D. (1978). Impersonal passives and the unaccusative hypothesis. In J. Jaeger (Eds.), *Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society* (pp.157-190). Berkeley, CA: Berkeley Linguistics society.

Pinker, S. (1987). The bootstrapping problem in language acquisition. In B. MacWhinney (Ed.), *Mechanisms of Language Acquisition* (pp.399-441). Hillsdale, NJ: Lawrence Erlbaum.

Rappaport Hovav, M., & Levin, B. (1998). Building Verb Meanings. *The Projection of Arguments: Lexical and Compositional Factors*, 97-134.

Rehurek, R., & Sojka, P. (2011). Gensim–python framework for vector space modelling. *NLP Centre, Faculty of Informatics, Masaryk University, Brno, Czech Republic*, 3(2).

Ritter, E., & Rosen, S. T. (1996). Strong and weak predicate: Reducing the lexical Burden. *Linguistic Analysis*, 26, 29-62.

Seabold, S., & Perktold, J. (2010). statsmodels: Econometric and statistical modeling with python. In *9th Python in Science Conference*.

Sorace, A. (2000). Gradients in auxiliary selection with intransitive verbs. *Language*, 76(4), 859–890. DOI: doi.org/10.2307/417202

Student. (1908). *The probable error of a mean*. Biometrika, 1–25.

Sun, J. (2012). Jieba Chinese word segmentation tool.

Wang, Arthur C.-A. (2010). *The Microparametric Syntax of Resultatives in Chinese Languages*. Doctoral dissertation, New York University.

Weber, J., & Colunga, E. (2022). How to Build a Toddler Lexical Network. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *44*. Retrieved from https://escholarship.org/uc/item/5hp3g2ss