Proceedings of the Second International Conference on Error-Driven Learning in Language (EDLL 2022)

Monday, 1 August — Wednesday, 3 August, 2022

University of Tübingen

Edited by Jessie S. Nixon, Fabian Tomaschek and R. Harald Baayen

Foreword

This volume contains the proceedings abstracts of the Second International Conference on Error-Driven Learning in Language, EDLL 2022, held online at the University of Tübingen, Germany, 1-3 August, 2022. The first conference of its kind was held last year, in March 2021.

We were inspired by the enthusiasm and energy participants brought to last year's conference, with the quality of the presentations, lively discussions and the warm, friendly and supportive atmosphere. So it is an honour and a pleasure to host this year's event, which, for the second time brings together researchers working in the exciting area of error-driven learning in language. We are also honoured to be joined by our excellent Keynote speakers: Jacolien van Rij, Padraic Monaghan, Elizabeth Wonnacott and Peter Dayan. Again, we have presenters and participants from around the world. Our programme includes a diverse range of experimental and computational modelling papers investigating different learning algorithms, as well as different aspects of linguistic processing and learning – from speech, morpho-phonetics and word learning to semantics, morphology, syntax and sentence processing.

Thanks to the ERC for financial support (grant number 742545). Most importantly, we would like to thank our Keynotes, presenters and all the participants for making this event possible.

Jessie Nixon
Fabian Tomaschek
Harald Baayen

Tübingen, July 2022

table of contents

Day 1
Monday, August 1, 2022

Day 2
Tuesday, August 2, 2022

13:30 - 14:00	Coffee/Tea	
	Keynote 2	
14:00 - 15:00	**Padraic Monaghan**	How multiple and multimodal sources of information support word learning
15:00 - 15:30	**Padraic Monaghan** Keynote session chair: Harald Baayen	Special Q&A for Junior Researchers
15:30 - 15:45	Tea/coffee break	
	Session 3	
15:45 - 16:10	**Erdin Mujezinović** & **Ruben van de Vijver**	To count or to predict? Revisiting error-driven distributional learning of phonetic categories
16:10 - 16:35	**Judit Fazekas,** **Andrew Jessop, Julian Pine** **& Caroline Rowland**	Do we learn from our prediction mistakes? Evaluating error-based theories of language acquisition
16:35 - 17:00	**Kaori Idemaru**, **Adam A. Bramlett** & **Vsevolod Kapatsinski**	Learning mechanisms in phonetic cue weighting: What do we learn from a single cue?
17:00 - 18:00	Poster session	
18:00 - 18:45	Lunch/dinner break	
	Keynote 3	
18:45 - 19:45	**Elizabeth Wonnacott**	Language input and generalization: Evidence from language learning experiments with adults and children
19:45 - 20:15	**Elizabeth Wonnacott** Keynote session chair: Michael Ramscar	Special Q&A for Junior Researchers

Poster session
Tuesday, August 2, 2022

17:00 - 18:00	
Kun Sun	Contextual semantic similarity computation using discriminative learning
Maria Heitmeier, **Yu-Ying Chuang** **& R. Harald Baayen**	Modelling trial-to-trial learning in lexical decision with Linear Discriminative Learning
Jessie S. Nixon, Sanne Poelstra **Jacolien van Rij**	When do we learn from absent cues?
Motoki Saito **& R. Harald Baayen**	The phonetic realization of German interfixes is co-determined by semantics

Day 3
Wednesday, August 3, 2022

9:30 - 10:00	Coffee/Tea	
	Session 4	
10:00 - 10:25	**Xuemei Chen, Suiping Wang & Robert Hartsuiker**	Structure prediction occurs when it is needed: Evidence from Visual-world Structural priming in Mandarin and Dutch comprehension
10:25 - 10:50	**Yanxin (Alice) Zhu & Theres Grüter**	Does structural priming lead to learning? Evidence from classroom learners of Mandarin
10:50 - 10:55	5-minute break	
10:55 - 11:20	**Giulia Bovolenta & Emma Marsden**	The effect of verb surprisal on the acquisition of new syntactic structures
11:20 - 11:45	**Judit Fazekas, Yamil Vidal, Julian Pine & Perrine Brusini**	The babe with the predictive power: work in progress examining the role of prediction in early word encoding
11:45 - 12:15	Break	
	Session 5	
12:15 - 12:40	**Simon David Stein & Ingo Plag**	Using linear discriminative learning to model the acoustic duration of English derived words
12:40 - 13:05	**Erdin Mujezinović Ruben van de Vijver**	Learning to unlearn: The role of negative evidence in morphophonological learning
13:05 - 13:30	**Dominic Schmitz, Viktoria Schneider & Janina Esser**	Evidence for a non-generic masculine generic in German
13:30 - 14:00	Break	
	Keynote 4	
14:00 - 15:00	**Peter Dayan**	Generative and discriminative reinforcement learning as model-based and model-free control
15:00 - 15:15	Closing remarks	
15:15 - 15:45	**Peter Dayan** Keynote session chair: Michael Ramscar	Special Q&A for Junior Researchers
15:45 ...	Social event	

— 1 —

keynotes

Cognitive modeling with error-driven learning

Jacolien van Rij
Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence,
University of Groningen
j.c.van.rij@rug.nl

In this talk, I will present three recent studies in which we used error-driven learning networks as *cognitive models* – i.e., computational simulations that aim to explain the cognitive processes involved in performing a task, such as reference processing. The first study investigated whether we can detect trial-by-trial learning in the ERP signal recorded during training of a pitch contrast. The second and third study investigate the roles of pre- and postmarking in learning noun classes.

These studies show that language processing is guided by expectations that follow from (discriminative) error-driven learning, but also raise questions for further research. For example, about how to simulate language processing and the interaction between error-driven learning with higher order cognition skills. To address some of these questions, we have started to use the cognitive architecture PRIMs in our projects that focus on the acquisition of sentence processing. I will present one of these studies, which aims to investigate how language learners acquire reference biases in language.

How multiple and multimodal sources of information support word learning

Padraic Monaghan
Department of Psychology, Lancaster University
p.monaghan@lancaster.ac.uk

Language learning has traditionally focused on information that derives from sequences of words in sentences, or sequences of sounds in words. However, this chronically underestimates the multifarious sources of information that are present in the environment both within the language and around the language (Bahrick et al., 2004). For instance, when hearing a word for the first time, children receive information about the phonemes in the word's sound, its context in a sentence, but also information from prosody in order to make the novel word salient, information from gesture to indicate the intended referent, as well as information from the objects, properties of objects, and actions that surround the child at the moment of naming (Yu et al., 2021). Understanding language learning, then, requires determining how the child navigates this buzzing teeming world of sound in situation.

In a series of computational modelling simulations paired with experimental studies, we have investigated how multiple sources of information may be useful for supporting the learner in determining how words map onto the world around them (Monaghan, 2017). In simulations and studies with adults, we found that learners can make use of prosody, gesture, and context of a word in supporting their learning of novel word-object mappings. In addition, we found that the substantial variability found in natural learning environments promoted learning more than if those multimodal information sources were perfectly reliable.

We also conducted studies with infants to determine whether similar patterns are observed in first language acquisition. In a looking time study, we found that 14 month old infants could use multiple information sources in guiding their learning (Dunn et al., in prep.). During exposure to novel word-object mappings, perfectly reliable information resulted in more looking to targets but less reliable (but still useful) information resulted in more looking time to the information sources rather than the target. During testing, reliable and less reliable information sources resulted in similar looking to targets. Caregivers, as well as children, are sensitive to multiple information sources in supporting learning. In a further study, we found that when asked to teach their child novel words when the environment contained multiple potential referents for the new word, caregivers produced more gestures than when the environment was unambiguous with respect to the target (Cheung et al., 2021).

The environment is noisy and variable, but we suggest that this noise is critical to robust, effective language learning, especially when caregivers contingently adapt their behaviour to support integration of multiple information sources in word learning.

References

Bahrick, L. E., Lickliter, R., & Flom, R. (2004). Intersensory redundancy guides the development of selective attention, perception, and cognition in infancy. *Current Directions in Psychological Science, 13,* 99-102.

Cheung, R. W., Hartley, C., & Monaghan, P. (2021). Caregivers use gesture contingently to support word learning. *Developmental Science*, e13098.

Dunn, K.J., Frost, R.C., & Monaghan, P. (in prep.). The effect of statistical frequency of multiple cues during infant cross-situational learning of word-referent mappings.

Monaghan, P. (2017). Canalization of language structure from environmental constraints: A computational model of word learning from multiple cues. *Topics in Cognitive Science, 9*, 21-34.

Yu, C., Zhang, Y., Slone, L.K., & Smith, L.B. (2021). The infant's view redefines the problem of referential uncertainty in early word learning. *Proceedings of the National Academy of Sciences of the United States of America, 118*, 52 e2107019118.

Language input and generalization: Evidence from language learning experiments with adults and children

Elizabeth Wonnacott
University of Oxford
elizabeth.wonnacott@education.ox.ac.uk

How do humans acquire their remarkable, productive linguistic capacities? Traditional approaches within linguistics have assumed this rests on innate knowledge of grammatical structures. My research explores the alternative: That simple but powerful learning processes are sufficient to detect and abstract consistent regularities in the linguistic environment. I seek evidence for this by conducting controlled learning experiments in which experimental participants are exposed to structures from a novel language, allowing us to precisely observe how learning relates to input distributions.

I will present data from two such learning experiments. Experiment 1 provides evidence from a foreign language training study with seven year olds (Viviani, Ramscar & Wonnacott, in prep)– that linguistic productivity arises from exposure to variable exemplars. I interpret this in terms of error-driven learning processes by which learners use prediction error and cue competition to dissociate irrelevant cues, and discriminate invariant relationships. Experiment 2 seeks more direct evidence of these learning processes. Specifically, we (Viviani, Ramscar & Wonnacott) have adapted a classic (adult) artificial language learning experiment (Ramscar et al. 2010) to incorporate real time eye-tracking measures during the exposure phase of the experiment. Data collection is ongoing, however eye-movements provide clear evidence that learners make fast predictions - and prediction errors - during learning. There are also preliminary indications that different patterns during training may link to different outcomes at test.

Generative and discriminative reinforcement learning as model-based and model-free control

Peter Dayan
Computational Neuroscience, Max Planck Institute for Biological Cybernetics
dayan@tue.mpg.de

Substantial recent work has explored multiple mechanisms of decision-making in humans and other animals. Functionally and anatomically distinct modules have been identified, and their individual properties have been examined using intricate behavioural and neural tools. One critical distinction, which is related to many popular psychological dichotomies, is between model-based or goal-directed control, which is reflective and depends on prospective reasoning, and model-free or habitual control, which is reflexive, and depends on retrospective learning. I will show how to see these two systems in generative and discriminative terms, respectively, and discuss their interaction and integration.

— 2 —

talk and poster sessions

Between two worlds: The case of Maltese morphology

Jessica Nieder[a] & Fabian Tomaschek [b]

nieder@phil.hhu.de, fabian.tomaschek@uni-tuebingen.de

[a]Department of General Linguistics, Heinrich-Heine-Universität Düsseldorf, Germany;
[b]Department of General Linguistics, Eberhard-Karls-Universität Tübingen, Germany

Maltese, a Semitic language spoken on the island country of Malta, has been described as a language that is situated in between two linguistic worlds: Romance, especially Sicilian Italian, vs. Semitic, especially Maghrebi Arabic varieties (Comrie, 2007). The situation of language contact due to Malta's geographical location between Sicily and the North-African coast led to a "split" or "hybrid" morphological system (Borg & Gatt, 2017; Mayer et al., 2013; Spagnol, 2011) that is, for example, visible in the Maltese plural formation: A variety of concatenative *sound* plurals, such as *omm-ommiijiet* 'mothers' contrasts with a variety of non-concatenative *broken* plurals, such as *kelb-klieb* 'dogs'

This study presents a computational modeling approach of this hybrid morphological system by using an error-driven learning algorithm (NDL) to classify Maltese nouns as belonging to the language families Semitic vs. Non-Semitic. NDL was trained on a data set of 2377 Tunisian nouns from Gugliotta and Dinarelli (2020) and 2377 randomly selected Italian nouns from the `subtlex-it` corpus[1]. The model was then tested on 6511 Maltese singular and plural nouns from Nieder et al. (2021).

The Maltese data set also included information about the origin of the nouns (Semitic vs. Non-Semitic) as well as information about number (singular vs. broken vs. sound). We used the same quasi-phonetic transcription (based on SAMPA) for all three languages. 2-phones were used as cues, the language family categories `semitic` vs. `non-semitic` were used as outcome.

The evaluation on the training data showed an excellent overall accuracy of 97.3%, indicating that NDL is able to capture the form differences between Semitic (Tunisian) and Non-Semitic (Italian) languages and to accurately classify nouns accordingly.

For the Maltese test data, we found that the network predicts the origin of the word forms with a high accuracy of 72%. The accuracies for predicted language family $\times$ number are given in Table 3 below.

Table 1: Origin: Semitic

	non-semitic	semitic
broken	16.54%	83.46%
singular	20.87%	79.13%
sound	15.2%	84.8%

Table 2: Origin: Non-Semitic

	non-semitic	semitic
broken	38.15%	61.85%
singular	54.63%	45.37%
sound	80.81%	19.19%

Table 3: NDL accuracies for nouns with a Semitic origin (left) vs. Non-Semitic origin (right), rows are the number categories, columns are the predicted language families `semitic` vs. `non-semitic`.

For nouns having a Semitic origin (Table 1), NDL provided accurate language family predictions ($\geq$79%) for all number categories. Interestingly, within the group of nouns with a Non-Semitic origin (Table 2), NDL was unsure how to classify singulars and broken plural nouns. This uncertainty is a direct reflection of the hybrid morphological system with nouns

[1]downloaded from http://crr.ugent.be/programs-data/subtitle-frequencies

such as *denfil* (sg.) - *dniefel* (broken pl.), etymologically related to Italian *delfino* 'dolphin', following a non-concatenative plural pattern despite having a Non-Semitic origin.

Our results show that it is possible to classify Maltese nouns using NDL with a training based on Tunisian and Italian nouns only. The model successfully applied the language-specific knowledge gained during the training to Maltese nouns, thus confirming the hybrid status of the Maltese morphological system.

References

Borg, C., & Gatt, A. (2017). Morphological analysis for the Maltese language: The challenges of a hybrid system. *Proceedings of the Third Arabic Natural Language Processing Workshop*, 25–34. https://doi.org/10.18653/v1/W17-1304

Comrie, B. (2007). Maltese and the World Atlas of Language Structures. *Introducing Maltese Linguistics: Selected Papers from the 1st International Conference on Maltese Linguistics* (pp. 3–15). John Benjamins Publishing.

Gugliotta, E., & Dinarelli, M. (2020). TArC: Incrementally and Semi-Automatically Collecting a Tunisian Arabish Corpus. *LREC*.

Mayer, T., Spagnol, M., & Schönhuber, F. (2013). Fixing the broken plural in Maltese. *Perspectives on Maltese Linguistics* (pp. 129–158). De Gruyter.

Nieder, J., Chuang, Y.-Y., van de Vijver, R., & Baayen, R. (2021). Comprehension, Production and Processing of Maltese Plurals in the Discriminative Lexicon. https://doi.org/https://doi.org/10.31234/osf.io/rkath

Spagnol, M. (2011). *A Tale of Two Morphologies: Verb structure and argument alternations in Maltese.* [Doctoral dissertation, University of Konstanz].

Simulating the learnability of Polish aspect: an error-driven approach

Irene Testini, Petar Milin & Dagmar Divjak
University of Birmingham

We take a learning-based approach to exploring the cognitive plausibility of the linguistic category of aspect in Polish. Slavic aspect is traditionally described as a binary category consisting of imperfective and perfective, with verbs forming imperfective/perfective pairs (e.g. *pisać_impf* versus *napisać_pf* - "write") or larger clusters, often triplets (e.g. *pisać_impf* versus *napisać_pf* versus *napisywać_secondary_imperfective*). In principle, both past and future tense allow the use of imperfective and perfective, with the present is limited to the imperfective. L2 learners are taught to choose one aspectual form, depending on the meaning they intend to express. Unfortunately, the meaning of aspect has not been easy to pin down, and more than half a dozen (abstract) semantic concepts (boundedness, totality, resultativeness, specificity, perspective, foregrounding and sequencing) have been proposed, much to the despair of the foreign learner (Janda 2003). To establish whether aspectual usage can be learned from exposure to the ambient language only, without the need for abstract semantic labels (e.g., boundedness, specificity), and to evaluate the types of cues and their informativity in learning to use the preferred aspectual form, we ran two studies, one in which we explore the natural distribution of aspectual forms, and then a learning simulation in which we learn aspect from exposure to usage.

We first focus on aspectual usage. We extracted 50,331,755 sentences in the indicative mood from the Araneum Polonicum corpus, excluding infinitives, participles, gerunds and agglutinatives. As the corpus is tagged using the National Corpus of Polish (NKJP) part-of-speech tagset only [https://www.sketchengine.eu/polish-nkjp-part-of-speech-tagset/], we wrote a set of heuristics to automatically annotate for tense and aspect from the available tags. Exploration of the raw data (Fig. 1) reveals that usage is biased towards the imperfective (70.7%) in general, the bulk being consumed by the present tense (57.0%). In the future, the perfective takes the lion share (9.2% versus 1.0% for the imperfective). Only in the past do we see a situation that could demand a choice to be made between imperfective and perfective, with imperfective (12.7%) and perfective (20.1%) past both occurring in decent numbers. However, tracking aspectual usage preferences for a set of 1,765 triplets revealed that the majority of verbs are biased towards one aspect, with only 282 verbs (16%) being equiprobable in either aspect, i.e., having a bias in the range of 0.4 – 0.6 (Fig. 3).

For the learning simulation, we employed the Naïve Discriminative Learning model (NDL; Baayen, Milin, Đurđević, Hendrix, & Marelli, 2011), which implements the Rescorla-Wagner rule of associative, error-correction learning. As NDL requires a frame of cues and outcomes for learning, we employed 3 different families of cues for each learning event (sentence): we extracted n-grams, groups of 1 to 4 contiguous words in a sentence to represent the context (similar to Romain et al. 2022); the verb's meaning (through so-called superlemmata that encompass the aspectual pair or triplet); and its tense (past, present or future). Filtering the corpus based on verbs for which we had superlemmata reduced it to 18,788,976 sentences; note that this reduction did not affect the original distribution of TA tags (Fig. 2). The model was then trained on a random stratified sample of 8,000,000 sentences and tested on 2,000,000. Table 1 displays F1 scores for each simulation. The results show that our algorithm achieves the highest accuracy (95% predicting imperfective and 90% predicting perfective) with a combination of superlemmata and basic tense information only.

Taken together, these results suggest the distribution of aspect usage is such that aspect is reliably predicted from a combination of lexical semantic and tense information alone: language users need to know what event is referred to and when it happened. Error-driven learning confirms that a high degree of accuracy in using aspect can be obtained without any linguistic abstractions related to aspect. For L1 speakers, aspect is thus likely a lexical phenomenon, learned on a lemma-by-lemma basis, taking into account whether the event has already taken place (past), is taking place (present) or will take place (future). For L2 learners, this means that the approach of offering aspectual pairs or triplets and a set of semantic rules to help choose the correct aspectual option may need to be retired: aspect would be more simply and more fruitfully thought of and taught by highlighting the limitations tense imposes.

Selected references

Baayen, R. H., Milin, P., Đurđević, D. F., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. Psychological Review, 118(3), 438-481.

Romain, L., Ez-zizi, A., Milin, P., & Divjak, D. (2022). What makes the past perfect and the future progressive? Experiential coordinates for a learnable, context-based model of tense and aspect. Cognitive Linguistics. https://doi.org/10.1515/cog-2021-0006.

Janda, L. 2003. A user-friendly conceptualization of Aspect". Slavic and East European Journal, vol 47, no. 2, 2003, pp. 251-281.

http://www.diaspol.uw.edu.pl/baza/?

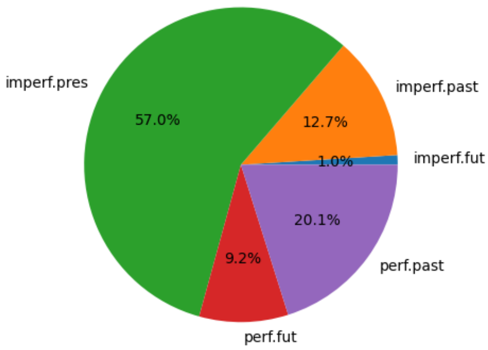


Fig. 1 Distribution of TA tags in full corpus

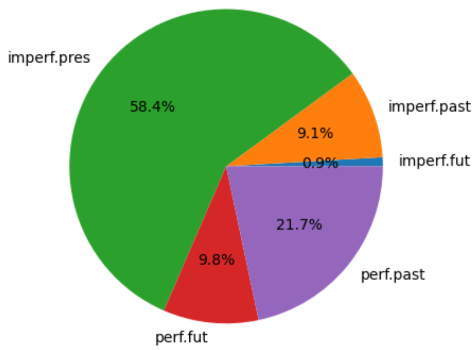


Fig. 2 Distribution of TA tags after filtering

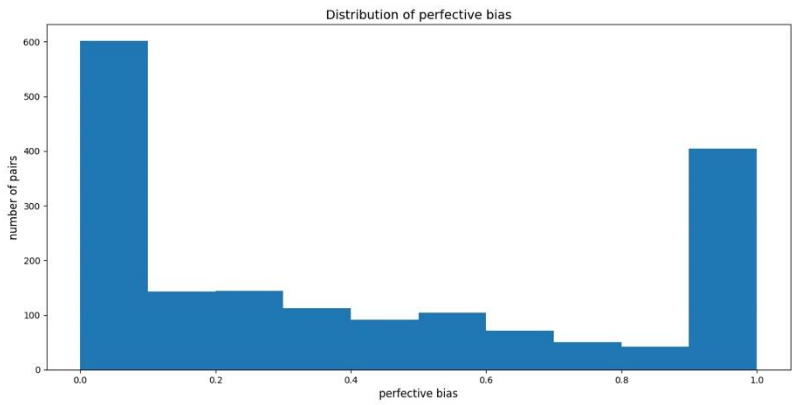


Fig. 3 Perfective Bias of Aspectual Pairs

CUES	F1 IMPERFECTIVE	F1 PERFECTIVE
ngrams	0.81	0.41
superlemmas	0.86	0.65
ngrams + lemmas	0.87	0.69
tenses	0.92	0.86
ngrams + tenses	0.93	0.86
superlemmas + tenses	0.95	0.9
superlemmas + ngrams + tenses	0.95	0.9

Table 1 Results of learning simulations

Distributed meanings and senses in error-driven learning framework – a proof of concept

Ksenija Mišić & Dušica Filipović Đurđević
The University of Belgrade, Faculty of Philosophy, Department of Psychology & Laboratory for Experimental Psychology
ksenija.misic@f.bg.ac.rs, dusica.djurdjevic@f.bg.ac.rs

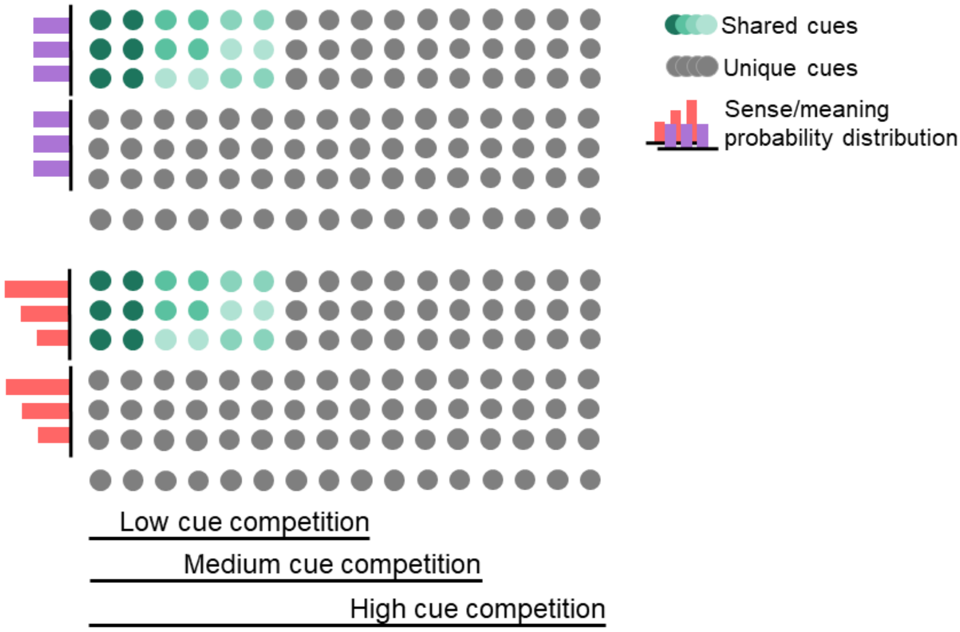
It was previously demonstrated that ambiguity effects could be simulated within error-driven learning framework (Filipović Đurđević & Kostić, 2021; Mišić & Filipović Đurđević, 2021). Filpović Đurđević and Kostić (2021) introduced a hypothesis that lexical ambiguity could be operationalized via partial overlap of multiple cues/outcomes related to meaning. Their demonstration relied on distributional semantics, namely the co-occurrence of words in the context. However, although relying on natural language samples is a powerful approach, it also introduces many complexities that potentially obscure the learning mechanics behind the ambiguity effects. Therefore, our aim was to perform ambiguity learning simulations from more of a theoretical standpoint by employing the toy model approach.

Theory informed our data generation process in two ways. First, error-driven learning (Rescorla & Wagner, 1972) offered a mechanism for learning ambiguous words and the importance of cue competition for learning to occur (Hoppe et al., 2022). Second, we relied on psycholinguistic theory for descriptions of ambiguity, namely the polysemy (multiple related senses) and homonymy (multiple unrelated meanings) distinction (Rodd et al., 2002). In addition to sense/meaning relatedness, we also paid attention to their probabilities (Filipović Đurđević & Kostić, 2017). We manipulated the type of lexical ambiguity (unambiguous words, polysemes, homonyms), the balance of sense/meaning probabilities (balanced, unbalanced), and the level of cue competition (low, medium, high).

Data were generated in the following way. We modelled a total of six words: two unambiguous words, a balanced and an unbalanced polyseme, and a balanced and an unbalanced homonym. Each word was represented by one outcome. Cues were created separately for each sense/meaning and were constructed as an equal-length string of arbitrary elements. The ambiguity type was manipulated via the cue overlap. Unambiguous words were predicted by a single cue set. Homonyms were predicted by three distinct sets of cues, each representing one meaning. Polysemous words were also predicted by three sets of cues, however, in addition to some unique cues for each of the senses, sets had some overlap among themselves in order to represent the sense relatedness. Each of the artificial words (outcomes) and its cues was presented to the network an equal amount of times. Balance of the sense/meanings frequency distribution was manipulated through the frequency of the presentation of each cue-outcome pairing. Finally, to introduce more cue competition, we randomly sampled a number of existing cues and appended them to other meanings/senses strings. By varying the number of cues appended, we varied the cue competition intensity. The data structure scheme is presented in Figure 1.

We then compared simulations on two different measures – the activation of the outcomes, and the learnability (a quantitative description of learning curves). When cue competition was present, activation decreased in the following order: balanced homonyms, unbalanced homonyms, unbalanced polysemes, and unbalanced polysemes, with unambiguous words, activated the least. This pattern, although expected to be inversely proportionate, was directly proportionate to the RTs in lexical decision tasks (Filipović Đurđević, 2019; Filipović Đurđević & Kostić, 2021). Learnability measure revealed that homonyms were learned the best, followed by polysemes, and then unambiguous words. Nevertheless, the existing relationship suggests that possible modifications of the generated data might lead to a better insight into how learning leads to the presence of ambiguity in language.

Figure 1. Artificial data structure



Filipović Đurđević, D. (2019). Balance of Meaning Probabilities in Processing of Serbian Homonymy. *Primenjena Psihologija*, *12*(3), 283–304. https://doi.org/10.19090/pp.2019.3.283-304

Filipović Đurđević, D., & Kostić, A. (2017). Number, Relative Frequency, Entropy, Redundancy, Familiarity, and Concreteness of Word Senses: Ratings for 150 Serbian Polysemous Nouns. In S. Halupka-Rešetar & S. Martinez-Ferreiro (Eds.), *Studies in Language and Mind* (pp. 13–77). Filozofski fakultet u Novom Sadu. http://digitalna.ff.uns.ac.rs/sadrzaj/2017/978-86-6065-446-7

Filipović Đurđević, D., & Kostić, A. (2021). We probably sense sense probabilitites. *Language, Cognition and Neuroscience*, 1–50. https://doi.org/https://doi.org/10.1080/23273798.2021.1909083 REGULAR

Hoppe, D. B., Hendriks, P., Ramscar, M., & Van Rij, J. (2022). An exploration of error-driven learning in simple two-layer networks from a discriminative learning perspective. *Behavior Research Methods*. https://doi.org/https://doi.org/10.3758/s13428-021-01711-5

Mišić, K., & Filipović Đurđević, D. (2021). Interaction of semantic and syntactic ambiguity in the light of discrimination learning. In J. S. Nixon, E. Shafaei-Bajestan, & R. H. Baayen (Eds.), *International Conference on Error-Driven Learning in Language (EDLL 2021) Book of Abstracts*.

Rescorla, R. A., & Wagner, A. R. (1972). A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: current research and theory* (pp. 64–99). Appleton-Century-Crofts. https://doi.org/10.1101/gr.110528.110

Rodd, J. M., Gaskell, G., & Marslen-Wilson, W. (2002). Making Sense of Semantic Ambiguity: Semantic Competition in Lexical Access. *Journal of Memory and Language*, *46*(2), 245–266. https://doi.org/10.1006/jmla.2001.2810

Delineating Prediction Mechanisms for Learning Words from Context

Layla Unger, Robert Ralston, Qianqian Wan, Mengcun Gao, and Olivera Savic
The Ohio State University

Early word learning may get off the ground by mapping words to referents that accompany them. For example, a child might learn the meaning of the word "apple" from hearing "apple" in the presence of apples. However, when a word occurs without a handy referent, the other words that accompany it might provide a rich source of information about its meaning, because *words similar in meaning occur in similar contexts of other words* [1–3]. For example, "apple" and "mango" both co-occur with words such as "juicy" and "sweet". Therefore, a child who has mapped "apple" to apples but never seen a mango may nonetheless learn that a mango is similar to an apple from their shared co-occurrence with "juicy" or "sweet". Thus, it is possible to learn new words by *linking them to known words that occur in similar contexts.* Such learning might be critical for extending word learning well beyond mapping individual words to referents. However, we know little about the learning mechanisms that support learning new words from context.

Proposals that language processing involves predicting upcoming words could provide candidate learning mechanisms for learning new words from context, because learning to predict upcoming words inherently involves learning regularities of word co-occurrence. However, to our knowledge, potential prediction-based mechanisms for word learning from context remain underspecified. Such proposals may allude to influential neural network models that learn to predict upcoming input [4], or present only a verbal description [5, 6]. Here, we evaluate two proposed mechanisms, which we refer to as Precursor Transformation and Co-Activation. Our aim is to highlight that these mechanisms are sensitive to different word context regularities that have different implications for word learning from context. Our future goal is to evaluate whether these mechanisms play important roles in word learning during child development.

In Precursor Transformation, input words are transformed into an internal representation that is used to predict the next input. Prediction errors update the *transformation of the preceding input* to generate more accurate predictions. This describes neural network models such as RNNs [4], which have formed the basis of influential models of prediction in language. For example, if "juicy" follows "apple", the mechanism will learn a transformation of "apple" that predicts "juicy". Here, we highlight that this mechanism only forms similar representations of words that predict similar upcoming words, or *share outcomes.* Thus, a new word should be learned more readily from context when it shares outcomes with known words. For example, a learner who has heard fruit words *preceding* "juicy" should readily learn "mango" when it also precedes "juicy". In contrast, words that *share predictors* should neither form similar representations, nor have a learning advantage.

The reverse pattern characterizes the Co-Activation mechanism. To our knowledge, this mechanism has only been described verbally [e.g., 5, 6]. According to such descriptions, the mechanism forms and modulates predictive links between words. For example, hearing "juicy apple" fosters a link between "juicy" and "apple" that subsequently causes "apple" to be predicted upon hearing "juicy". Critically, prediction errors also *foster links between a co-active erroneous word and an observed word.* For example, when the learner hears "juicy" and predicts "apple" but instead hears "mango", "apple" and "mango" become linked. Thus, the Co-Activation mechanism should only be sensitive to the regularity with which words *share predictors.*

The distinction between sensitivity to shared outcome versus shared predictor regularities is meaningful because different regularities may characterize different sets of words similar in meaning. For example, words similar in meaning that tend to occur early in utterances (e.g., words for animate agents) may primarily share outcomes, whereas those that tend to occur later (e.g., words for inanimate objects) may primarily share predictors. The present work demonstrates these qualitative distinctions by implementing these two mechanisms, and simulating them with toy languages that differentiate shared outcome and shared predictor regularities. We used the Keras API to (1) implement Precursor Transformation as an RNN, and (2) create a novel implementation of Co-Activation. In addition, we developed a novel approach for probing new word learning from context. Our simulations confirm that new word learning depends on shared outcome regularities in the Precursor Transformation mechanism, and shared predictor regularites in the Co-Activation mechanism.

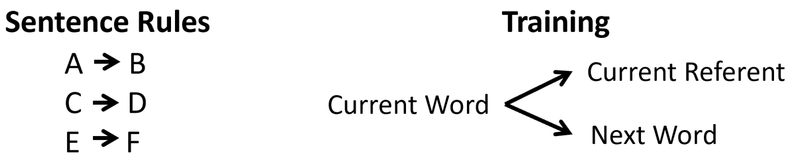


Figure 1. Toy language training. Each letter is a "class" of multiple words that occur in the same sentence contexts (e.g., A-words always precede B-words). Sentences were randomly concatenated for training. Each word also had a corresponding referent. Words and referents were orthogonal one-hot vectors. Models were trained to map words to their referents and predict the next word.

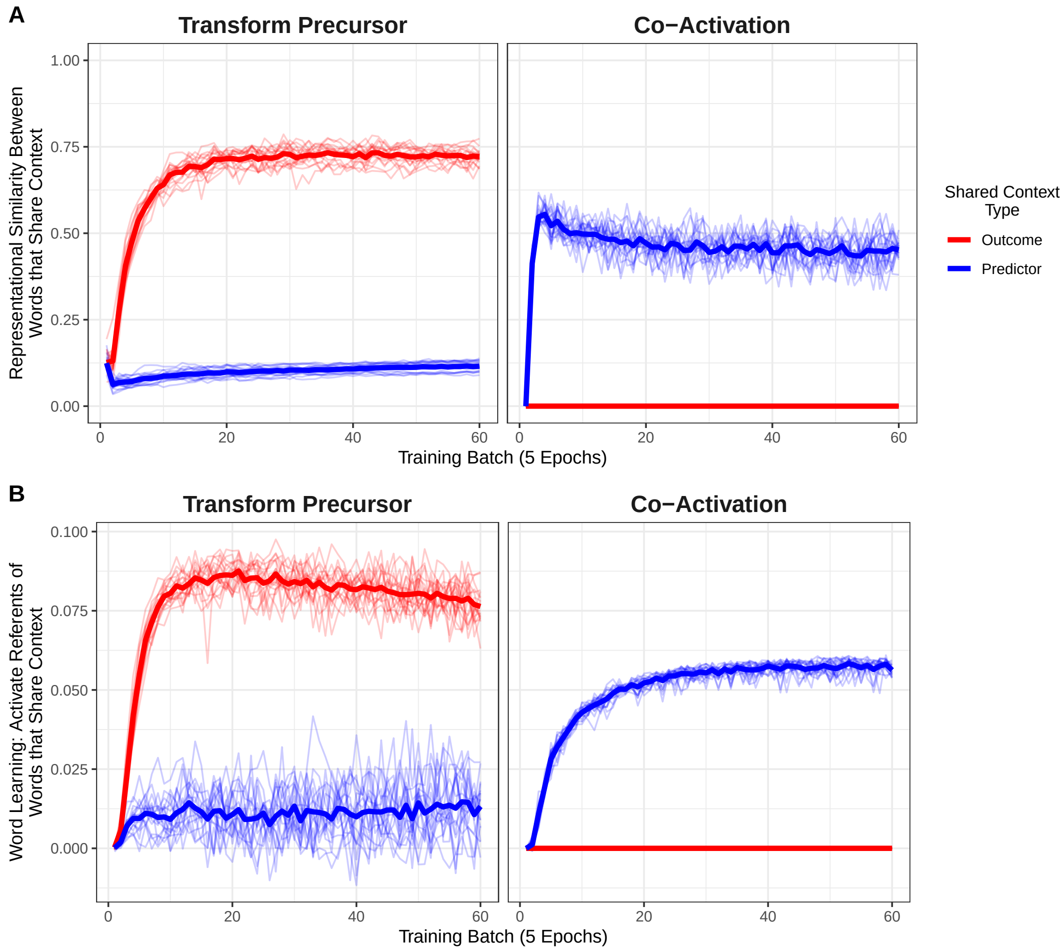


Figure 2. Models were simulated 20 times and probed every 5 epochs to assess representational similarity and new word learning. A. Representational similarity was calculated as the similarity between words that share context versus words that occur in the same sentence position but different contexts. B. Word learning was probed by giving the model 5 epochs of training on sentences containing one new word *without* referents, and assessing how well it mapped the new word to referents of words that occurred in the same context.

[1] Landauer, T.K. and Dumais, S.T. (1997). *Psychological Review*, 104(2), 211–240.
[2] Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013). *arXiv:1301.3781*.
[3] Miller, G.A. and Charles, W.G. (1991). *Language and Cognitive Processes*, 6(1), 1–28.
[4] Elman, J.L. (1990). *Cognitive Science*, 14(2), 179–211.
[5] Ervin, S.M. (1961). *The American Journal of Psychology*, 74(3), 361–372.
[6] McNeill, D. (1963). *Journal of Verbal Learning and Verbal Behavior*, 2(3), 250–262.

The logistic perceptron accounts for rank frequency effects in lexical processing

Vsevolod Kapatsinski (University of Oregon)

Murray and Forster (2004) showed that rank frequency is a better linear predictor of lexical decision times than log frequency (replicated in Figures 1 and 2 based on frequencies of verbs and lexical decision times from Balota et al., 2007). Murray and Forster argued that this finding contradicts parallel models of processing, and supports Forster's (1976) serial search model of word recognition. Specifically, the rank frequency of a word is the number of steps it would take to find the word by going down a list of words ordered in inverse frequency order. This idea has recently risen in prominence again because Yang (2016) used it to derive a theory of productivity: according to Yang, a rule is productive if it wouldn't take too long to serially search the list of exceptions to the rule. In this talk, I show that Murray and Forster's results do not imply serial search, and are fully consistent with parallel processing.

Indeed, error-driven learning models that learn connections from sublexical cues to a word suggest that the activation reaching a word node should be roughly proportional to its log frequency (Baayen, 2010; Olejarczuk et al., 2018). However, this does not mean that lexical decision times should be a linear function of this activation. First, parallel models make categorical decisions (such as lexical decision) by activating nodes with logistic activation functions (see Plaut & Booth, 2000, for lexical decision specifically). These nodes sum inputs and transform the sum through an S-shaped function bounded between 0 and 1, as probabilities must lie within this interval. Second, rank frequency turns out to be an S-shaped function of log frequency. Therefore, Murray and Forster's results are consistent with the lexical decision time tracking the output activation of a decision node whose input activation is proportional to the log frequency of the word (i.e., a logistic transformation of log frequency, as shown in Figure 3).

The superiority of rank frequency over log frequency in predicting lexical decision times is fully consistent with standard assumptions about how processing and learning work in parallel models. It provides no evidence for serial search.

References:

Baayen, R. H. (2010). Demythologizing the word frequency effect: A discriminative learning perspective. *The Mental Lexicon*, *5*(3), 436-461.

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., ... & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*(3), 445-459.

Forster, K. I. (1976). Accessing the mental lexicon. In R. J. Wales & E. Walker (Eds.), *New approaches to language mechanisms* (pp.257-287). Amsterdam: North Holland.

Murray, W. S., & Forster, K. I. (2004). Serial mechanisms in lexical access: the rank hypothesis. *Psychological Review*, *111*(3), 721-756.

Olejarczuk, P., Kapatsinski, V., & Baayen, R. H. (2018). Distributional learning is error-driven: The role of surprise in the acquisition of phonetic categories. *Linguistics Vanguard*, *4*(s2).

Plaut, D. C., & Booth, J. R. (2000). Individual and developmental differences in semantic priming: empirical and computational support for a single-mechanism account of lexical processing. *Psychological Review*, *107*(4), 786-823.

Yang, C. (2016). *The price of productivity: How children learn to break the rules of language*. MIT Press.

Figure 1. Lexical decision times for verbs from the English Lexicon Project (Balota et al., 2007) as a function of log frequency

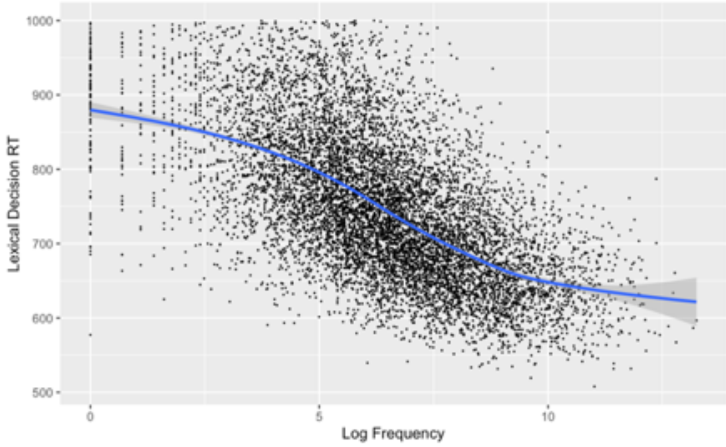


Figure 2. Lexical decision times for verbs from the English Lexicon Project (Balota et al., 2007) as a function of rank frequency

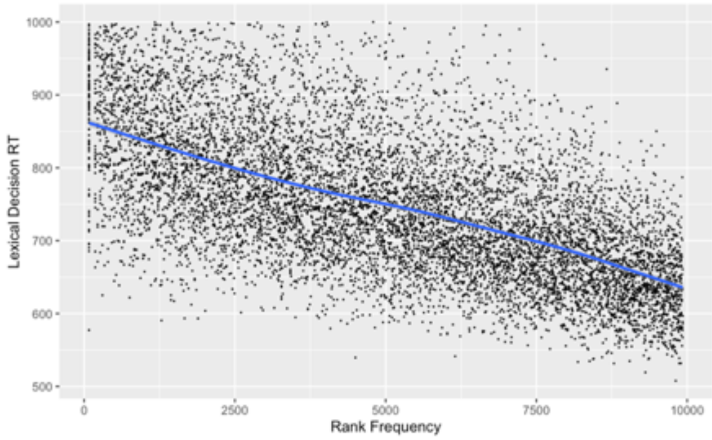
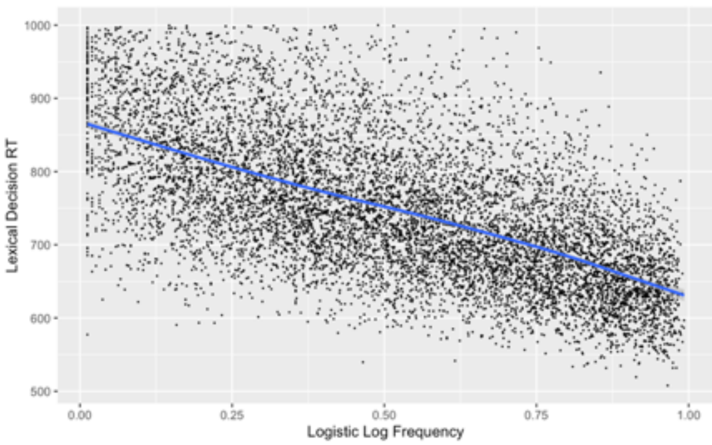


Figure 3. Lexical decision times for verbs from the English Lexicon Project (Balota et al., 2007) as a function of logistic log frequency



Syntagmatic paradigms: Temporal proximity increases salience of paradigmatic cues

Vsevolod Kapatsinski and Amy Smolek (University of Oregon)

Speakers of morphologically rich languages often face the need to produce novel forms of words that they know. In doing so, they appear to rely both on the meaning they need to express (semantic cues) and on knowledge of other forms of the same word (paradigmatic cues). In early work on paradigm learning, it was suggested that paradigmatic mappings lik CCoC→CCCa could not be learned through predictive error-driven learning because the forms to be associated are never next to each other (McNeill, 1966). However, more recently, corpus studies have shown that paradigmatically-related words in fact do frequently co-occur, to the point that co-occurrence can be used as a sign of paradigmatic relatedness by computational models (Baroni et al., 2004; Xu & Croft, 1998). These findings motivated our hypothesis that the instances of temporal proximity, which involve paradigmatically-related forms being close to each other in time, are crucial for learners to acquire paradigmatic mappings. We examine how temporal proximity influences the acquisition of paradigmatic mappings in human learners, and show that these effects of proximity are captured by differences in salience of paradigmatic cues in the Rescorla-Wagner model (Rescorla & Wagner, 1972).

Specifically, native English speakers were exposed to miniature artificial languages with two plural suffixes, [i] and [a], in which [a] triggered a change in the preceding consonant, turning either the velars, [k] and [g], (for some subjects) or the labials, [p] and [b], (for other subjects) into alveopalatals, [tʃ] and [dʒ]. Participants were assigned to one of the following 4 conditions, crossing whether corresponding forms featuring a stem change (e.g., k→tʃa) were temporally proximate, and whether corresponding forms not featuring a stem change (e.g., t→ta, t→ti) were temporally proximate (SG=Singular, PL=plural).

(1) All Obvious: ***blup*$_{\mathbf{SG}}$ *blupa*$_{\mathbf{PL}}$** *klut*$_{\mathrm{SG}}$ *kluti*$_{\mathrm{PL}}$ ***smak*$_{\mathbf{SG}}$ *smatʃa*$_{\mathbf{PL}}$**…
(2) Only Change Obvious: *kluti*$_{\mathrm{PL}}$ *blup*$_{\mathrm{SG}}$ *klut*$_{\mathrm{SG}}$ ***smak*$_{\mathbf{SG}}$ *smatʃa*$_{\mathbf{PL}}$** *blupa*$_{\mathrm{PL}}$…
(3) Only NoChange Obvious: ***blup*$_{\mathbf{SG}}$ *blupa*$_{\mathbf{PL}}$** *smatʃa*$_{\mathrm{PL}}$ ***klut*$_{\mathbf{SG}}$ *kluti*$_{\mathbf{PL}}$** *smak*$_{\mathrm{SG}}$…
(4) Neither Obvious: *kluti*$_{\mathrm{PL}}$ *blup*$_{\mathrm{SG}}$ *smatʃa*$_{\mathrm{PL}}$ *klut*$_{\mathrm{SG}}$ *smak*$_{\mathrm{SG}}$ *blupa*$_{\mathrm{PL}}$…

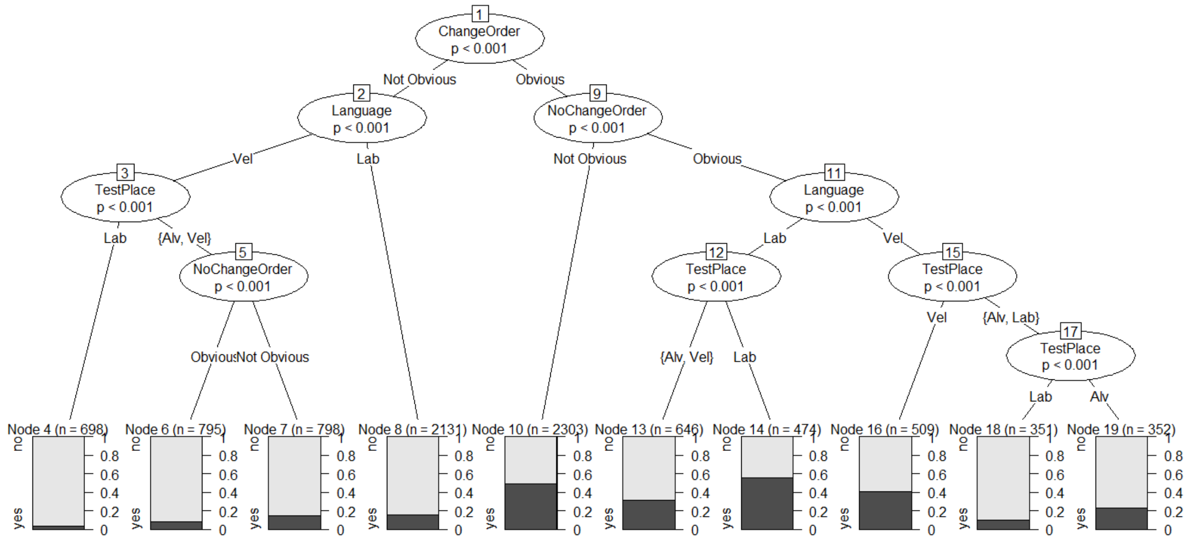
Knowledge of the language was then tested by presenting participants with novel singulars and corresponding plural meanings, and asking them to say the plural form. We found that temporal proximity of forms exemplifying a paradigmatic mapping increased the use of that mapping (Table 1). Specifically, the Change of singular-final consonants into [tʃa] was most productive in the Only Change Obvious condition, followed by All Obvious, followed by Neither Obvious, followed by Only NoChange Obvious. Furthermore, only participants in the All Obvious condition learned what consonants should be changed. Those in the Change Obvious condition instead changed alveolars, [t] and [d], as much as the consonants they heard change (Fig. 1).

We then modeled what participants learned using the predictive error-driven Rescorla-Wagner model, with semantic cues like SG vs. PL, syntagmatic cues (suffix vowel), and paradigmatic cues like Place of articulation of the final consonant of the singular form predicting the shape of the plural form's stem. To implement out hypothesis that proximity influences salience, we fit cue and outcome salience parameters of the model separately for each condition, minimizing the absolute deviation of the observed and predicted choice probabilities for each combination of Language, suffix and stem-final consonant (MAD). As expected, syntagmatic, within-form associations are unaffected by form proximity. However, crucially, temporal proximity between corresponding forms increases the salience of paradigmatic cues exemplified by those forms, as well as the salience of outcomes that indicate the differences between them (Table 2). The observed changes in salience provide a possible mechanism for learning paradigmatic mappings using the same error-driven learning mechanisms used for learning syntagmatic associations.

Table 1. Change rates as a function of whether Change and NoChange were obvious in a logistic mixed effects model fit with lme4 (Bates et al., 2015). No further interactions were significant. Maximal random-effects structure was used. TBC = singular-final consonants that were to be changed according to the exposure language.

	b	*se(b)*	*z*	*p*
(Intercept)	-3.0428	0.3954	-7.695	<.00001
TBP = no	0.134	0.1377	0.973	.33
Change Obvious	3.4397	0.462	7.445	<.00001
Change Obvious x TBC = no	-0.7212	0.147	-4.907	<.00001
NoChange Obvious	-0.8616	0.4515	-1.908	.056
NoChangeObvious x TBC = no	-0.5945	0.1401	-4.242	.00002

Figure 1. Conditional inference tree showing that only the participants in the All Obvious Condition (ChangeOrder = Obvious & NoChangeOrder = Obvious) learned the effect of consonant place of articulation they were exposed to (TestPlace), changing Lab(ials) more than other consonants in the Lab(ial) language and Vel(ars) more than others in the Vel(ar) language.



To count or to predict? Revisiting error-driven distributional learning of phonetic categories

Erdin Mujezinović & Ruben van de Vijver
Heinrich-Heine University Düsseldorf

Olejarczuk, Kapatsinski, and Baayen (2018) provided striking evidence that it is not raw counts that determine the categorical representations, but log-odds, despite the fact that learning within linguistics is often viewed as statistical and assumes the acquisition of phonetic categories through faithful statistical tracking based on positive evidence only (Maye, Werker, and Gerken, 2002, Saffran, 2020, Saffran and Kirkham, 2018). In statistical learning a learner tracks the modality of phonetic distributions through an unsupervised cluster-analysis: If a phonetic parameter clusters around one value (is unimodal), there is only one category; if the parameter clusters at two different values (is bimodal), two categories emerge. These categories are either based on the mean or the median of the distribution.

Discriminative error-driven learning models (Hoppe et al., 2022, Rescorla and Wagner, 1972) attribute the influence of log-odds, as opposed to raw counts, in learning to prediction and prediction error. In phonetic category learning listeners try to predict subsequent tokens (outcomes) through previously experienced tokens (cues). As infrequent tokens are less expected, the representations shift towards the distribution's tail. Rare category member are more surprising and generate stronger prediction error, which in turn results in stronger category updating. Olejarczuk, Kapatsinski, and Baayen (2018) have confirmed such mechanisms to be present during the learning of a rise-fall tonal category ([ka]-syllables superimposed with a L-H-L tonal pattern).

As tones are not lexical in English, it may be the case that the effect found by Olejarczuk, Kapatsinski, and Baayen (2018) is limited to non-lexical categories. We therefore conducted a pilot study with 14 German native speakers which were taught a length distinction in an existing lexical category: lengthened /lː/. Two groups, *right-skew* (color-coded in blue) and *left-skew* (color-coded in red) listened to 256 [alːa]-tokens. The duration of the [lː]-tokens were manipulated in 20 ms increments. The tail of the right-skew group had higher, the tail of the left-skew group had lower durations. The differences between both groups are summarized in Table 1. Participants had to judge the category-goodness for each duration-step using a 7-point Likert scale. After rating the category-goodness, they were additionally asked to produce 3 representative [alːa]-tokens. The results are shown in Fig. 1.

Both category-goodness and production data corroborate the findings in Olejarczuk, Kapatsinski, and Baayen (2018), with the shape of the ratings being shifting towards the tails (lower durations = worse for right-skew, higher durations = worse for left-skew). The representative [lː]-durations were lower for the left-skew group than for the right-skew group. We thus have successfully replicated and extended the findings in Olejarczuk, Kapatsinski, and Baayen (2018) for consonant duration. We have also shown that the disproportional effect of infrequent items is observable even when the exposure phase was reduced (256, compared to 512 trials in Olejarczuk, Kapatsinski, and Baayen (2018)). Our findings support the notion of distributional learning being based on prediction and prediction-error, not on raw counting of the statistics.

Group	Range	Modal	Mean	Tokens
Right-skew	160-420 ms	220 ms	280 ms	256
Left-skew	140-400 ms	340 ms	280 ms	256
Test	80-480 ms			21

Table 1: Difference in distribution between right-skew and left-skew

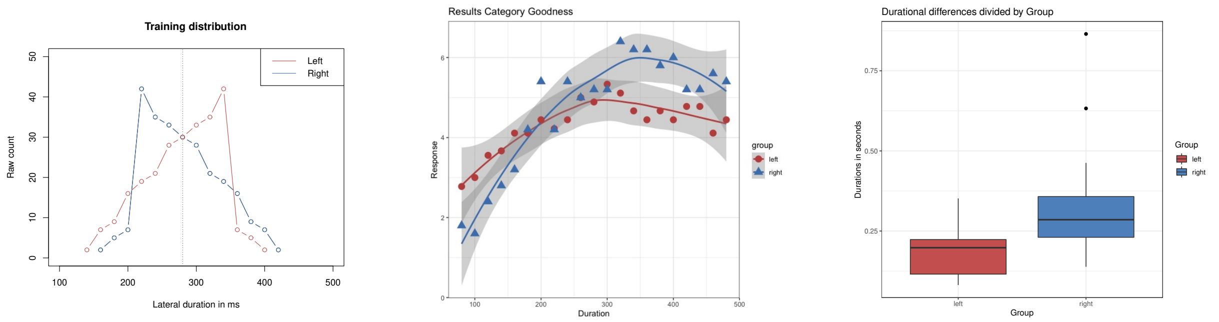


Figure 1: Training distributions (left), results "category-goodness" (middle) & results "production study" (right)

References

Hoppe, D. B., Hendriks, P., Ramscar, M., & van Rij, J. (2022). An exploration of error-driven learning in simple two-layer networks from a discriminative learning perspective. *Behavior Research Methods*, 1–31.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3), B101–B111.

Olejarczuk, P., Kapatsinski, V., & Baayen, R. H. (2018). Distributional learning is error-driven: The role of surprise in the acquisition of phonetic categories. *Linguistics Vanguard*, *4*(s2).

Rescorla, R. A., & Wagner, A. R. (1972). A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement (A. H. Black & W. Prokasy, Eds.). *Classical conditioning II: current research and theory*, 64–99.

Saffran, J. R. (2020). Statistical language learning in infancy. *Child Development Perspectives*, *14*(1), 49–54.

Saffran, J. R., & Kirkham, N. Z. (2018). Infant statistical learning. *Annual review of psychology*, *69*, 181–203.

Do we learn from our prediction mistakes? Evaluating error-based theories of language acquisition

Judit Fazekas[a], Andrew Jessop[b], Julian Pine[b] and Caroline Rowland[c]
[a]School of Psychological Sciences, University of Manchester, Ellen Wilkinson Building, Manchester M13 9PL, UK
[b]Department of Psychological Sciences, University of Liverpool Institute of Psychology Health and Society, Liverpool, UK
[c]Language Development Department, Max Planck Institute for Psycholinguistics, Nijmegen, Gelderland, The Netherlands

The role of prediction in language processing has been thoroughly examined. But while its role in language acquisition is frequently discussed, it has been much less extensively tested. Error-based theories of language acquisition (e.g. Chang et.al., 2006) propose that children and adults constantly predict the next upcoming word, compare these predictions to the actual input and, if there is a discrepancy, produce an error-signal which then prompts learning. Thus, a strong – but scarcely tested – prediction of these theories is that surprising input leads to more learning than predictable input.

In our study, we used the prime surprisal paradigm which has previously demonstrated that sentences in which the dative structure is more surprising (because the verb more frequently appears with an alternative structure) leads to stronger priming than predictable sentences (e.g. Peter et al., 2015). We embedded this paradigm in a four-phase intervention study to test this hypothesis. 72 monolingual English-speaking adults and 72 5-6 year-old children participated in the experiment. They took turns with the experimenter to describe video animations depicting transitive actions that could be described using either prepositional (PD) or double object dative (DOD) sentences. We report the data from the first 3 phases targeting abstract learning effects.

Phase 1 – Baseline phase: Measured participants' baseline dative production while the experimenter described filler videos.
Phase 2 – Bias or test-phase: The main goal of this phase was to induce long-term abstract learning effects where we can differentiate between learning rates for predictable and surprising sentences. We split the participants into two equal groups: the DOD-bias group always heard sentences where the DOD structure was unexpected and the PD structure predictable (given the verb's bias). The PD-bias group experienced the opposite pattern. Both groups heard the same overall number of DOD and PD structures. Our second goal was to replicate the immediate prime surprisal effects found in earlier studies.
Phase 3 – Post-test: Re-measured participants' post-bias dative production to assess whether participants pre- to post-test production moved towards the dative structure they were biased towards.

Results:
Phase 1: Adults showed an average 65.7% and children an average 31.3% baseline DOD production.
Phase 2, *Figure 1.:* The adult, but not the child group showed a significant prime surprisal effect when baseline performance was taken into account.
Phase 3, *Figure 2:* Both adults' (average 4.25% shift) and children's (average 6% shift) pre- to post-test production shifted towards the dative structure they were exposed to in surprising sentences. This effect was significant in the two age groups together and in the child group separately when excluding participants who showed ceiling performance at the pre-test phase.

In conclusion, our results support the central claim of error-based learning theories that surprising input leads to more learning than predictable input and the numerically larger child results potentially indicate life-long learning at a decreasing rate. The longer-term effects (phase 3) in the absence of immediate prime surprisal (phase 2) in the child group raises questions about whether immediate prime surprisal effects and long-term learning are both induced by the same mechanism.

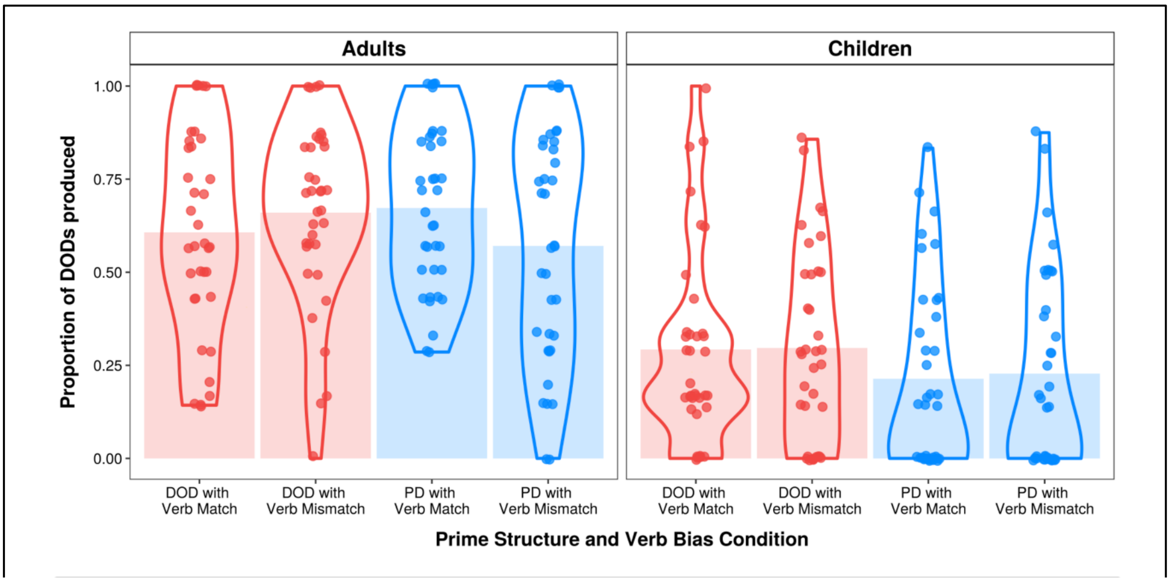


Figure 1. – Proportion of DOD production in the bias phase per age group and condition.

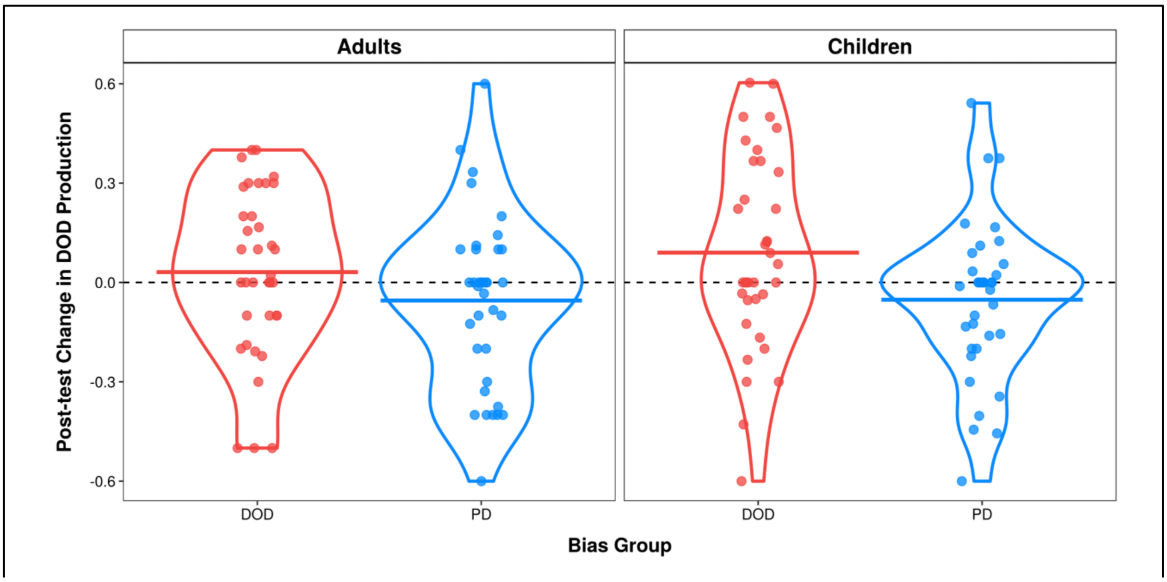


Figure 2. – Pre-to post-test difference per age group and bias group. The dashed line represents no pre- to post-test change while the solid lines show the average per age- and bias-group shifts.

References:

Chang, F., Dell, G. S., Bock, K. (2006). Becoming syntactic. Psychological review, 113, 234. (doi:10.1037/0033-295X.89.1.1)

Peter, M., Chang, F., Pine, J. M., Blything, R., Rowland, C. F. (2015). When and how do children develop knowledge of verb argument structure? Evidence from verb bias effects in a structural priming task. Journal of Memory and Language, 81, 1-15.

Learning mechanisms in phonetic cue weighting: What do we learn from a single cue?

Kaori Idemaru (University of Oregon), Adam Bramlett (Carnegie Mellon University), and Vsevolod Kapatsinski (University of Oregon)

Phonetic cue weighting refers to weighting acoustic/phonological dimensions like voice onset time (VOT) and fundamental frequency (F0) at the beginning of the following vowel as cues to phonological categories like /p/ or /b/ or lexical/semantic categories like PIER and BEER. Cue reweighting is frequently needed to adapt to a second language or an unfamiliar accent. For example, native English speakers listening to Korean-accented English must upweight F0 and/or downweight VOT to perceive Korean-accented speech accurately because Koreans rely primarily on F0 in differentiating English stops in production.

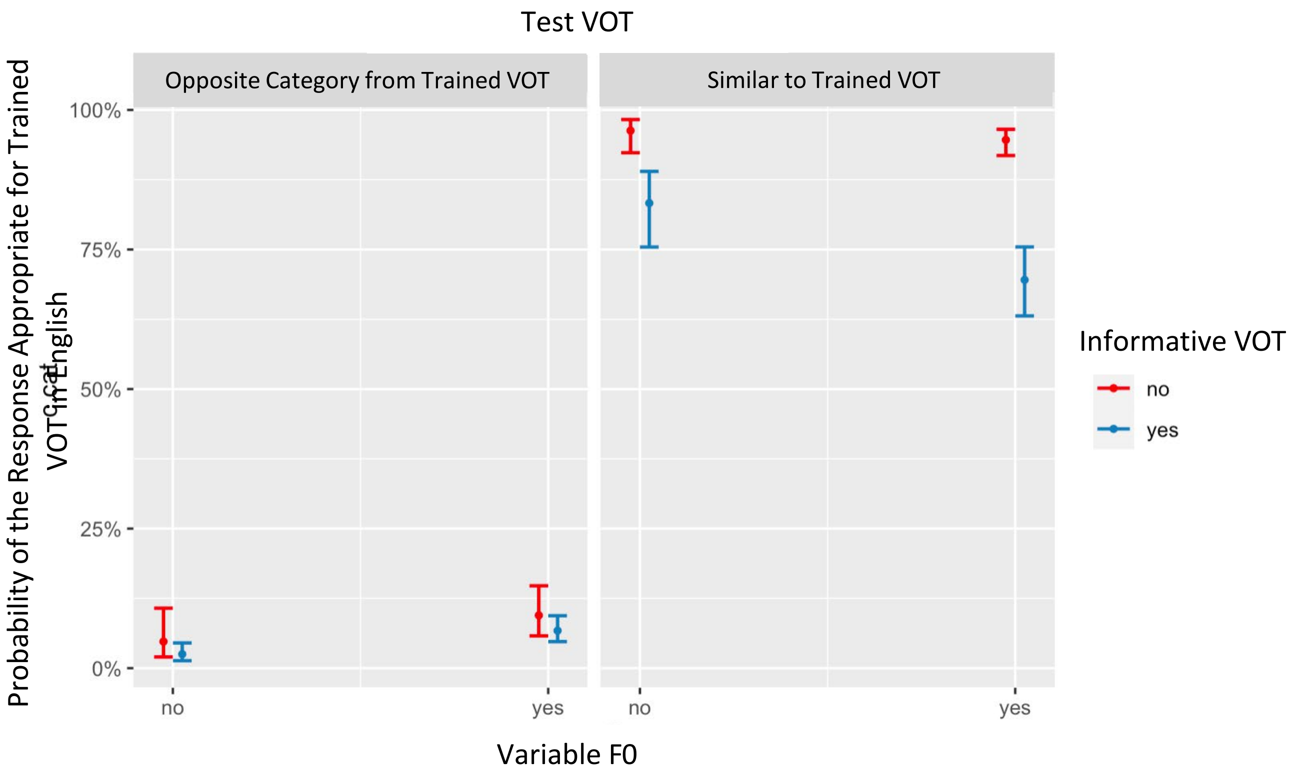
Harmon et al. (2019; HIK) tested some predictions of the Rescorla-Wagner model (RW; Rescorla & Wagner, 1972) for this process. They exposed English speakers to an artificial accent in which VOT was variable and not predictive of voicing/lexical identity whereas F0 was either predictive or constant. The Rescorla-Wagner model predicts downweighting of VOT in both conditions. In contrast to this prediction, the participants downweighted VOT only when they had another cue to rely on in training, i.e., when F0 was variable and predictive of the correct response. Based on these results, they proposed a new model of cue weighting. Specifically, they argued that learners reallocate attention across perceptual dimensions like VOT and F0 when doing so would reduce prediction error or maximize reinforcement.

The present experiment tests these models' predictions by exposing learners to only a single value of the primary cue dimension (VOT). This is interesting because the HIK model predicts that listeners can reweight entire dimensions based on a single experienced value, e.g., generalizing from short VOT being uninformative to all values of VOT being uninformative.

Native English listeners (N=221) participated in a perceptual training experiment in which they heard words like pier and beer, and were asked to indicate what they heard. In the training stimuli, VOT was held constant at the value of 5 ms, 25 ms or 45 ms, while F0 was either informative or held constant at 215 Hz, an ambiguous value. When it varied, it was predictive of feedback, cueing /b/ at 180 Hz, and /p/ at 250 Hz. At test, participants responded to VOT = 15 or 35 ms and F0 = 183 and 233 Hz (values half-way between the uninformative and informative values used in training). A pretest/posttest design was used to assess pre-training beliefs and their change with training. Prior to training, the participants believed that VOT = 45 ms is /p/, VOT = 5 ms is /b/, and VOT = 25 ms is ambiguous, consistent with the long-term English pattern. However, in all training conditions, the experienced value of VOT was paired with /p/ and /b/ feedback half the time. Thus, the /p/ feedback in the VOT = 5 ms condition is unexpected, as is /b/ feedback when VOT = 45 ms. In contrast, /p/ and /b/ are expected to occur about 50% of the time in the VOT = 25 ms condition. Therefore, the VOT=5 and VOT=45 conditions are expected to produce more prediction error, resulting in more learning.

Both HIK and RW predict that there should be more learning when VOT is informative (i.e., VOT=5 and 45). Figure 1 shows that this prediction was confirmed. In addition, HIK predicts that reliance on the entire VOT dimension would decrease, and only when F0 is predictive of the correct response. In contrast, RW predicts that the experienced VOT value would be reassociated, and more when F0 cannot explain the changes in voicing. As Figure 1 shows, the results are intermediate between the two sets of predictions. As HIK predicts, there is more learning about VOT when F0 is variable. However, some learning still occurs when F0 is constant. Furthermore, participants did not appear to downweight VOT as a *dimension*: although responding with the unexpected outcome (e.g., /p/ for VOT = 15) increased for VOT values similar to the experienced one, this learning did not generalize fully to VOT values in the other voicing category (e.g., learning that VOT = 35 is an unreliable cue to [p] from observing that VOT = 5 is an unreliable cue to [b]). This is consistent with reassociating experienced VOT values, as in RW, rather than withdrawing attention from the VOT dimension. Learners are reluctant to downweigh cues when they have nothing else to rely on (HIK), but they appear to reassociate experienced cues (RW) rather than (or in addition to) downweighting dimensions.

Figure 1. Predictions of a Bayesian logistic regression model fit to the human data in brms (Bürkner, 2017) for the posttest. Left panel: The test VOT value in the opposite voicing category from the trained VOT value (VOT = 15 for training with VOT = 45; VOT = 35 for training with VOT = 5). Right panel: The test value within the trained voicing category (VOT = 35 for training on VOT = 45; VOT = 15 for training on VOT = 5). Informative VOT is whether participants were trained on informative (5 or 45, "yes") or uninformative (25, "no") VOT values. The vertical axis is the probability of responding with the outcome that was predicted erroneously during training, i.e., the one participants expected based on their prior, English experience ([b] for those trained on VOT=5, [p] for those trained on VOT=45, whichever was more likely for VOT=25 at pretest for those trained on VOT=25). Learning means decreasing responding with this outcome for the trained VOT. Dimensional reweighting is observed if responding with the same outcome increases for the VOT from the other category. Learning is observed in both Informative VOT conditions (blue) for the VOT value within the trained category (right panel), but more learning is observed when F0 is variable (Variable f0 = yes).



References:
Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*, 1-28.
Harmon, Z., Idemaru, K., & Kapatsinski, V. (2019). Learning mechanisms in cue reweighting. *Cognition*, *189*, 76-88.
Rescorla, R. A, & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory.* Appleton Century Crofts.

Contextual semantic similarity computation using discriminative learning

Kun Sun
Department of Linguistics, University of Tübingen

Semantic similarity computation which has been investigated for decades is a key component of various applications, such as computational linguistics, cognitive psychology, and artificial intelligence. Currently semantic similarity computation is usually based on word vectors. Word vectors yielded by the discriminative learning model (e.g., word-to-word, word-to-ngram) are interpretable, compared with the state-of-the-art word embeddings trained by some popular programs (e.g. word2vec, BERT). However, semantic similarity computed based on these popular word vectors only capture some semantic features such as distributional statistics and co-occurrences. Any numbers in such a word vector is not interpretable. By contrast, word vectors trained by discriminative learning can form a matrix of cue weights, and the matrix indicates the degree to which each word in the trained corpus predicts each other word in this corpus. In this sense, any number within one word vector trained by discriminative learning represents the semantic association between two words, and the semantic association is actually one sort of semantic activation, which is closely related with word meaning. Word vectors trained by discriminative learning can capture semantic activation, which is quite distinct from the past paradigms of word vectors.
Although loads of studies have proposed algorithms to assess semantic similarity, semantic similarity was computed between two words without a context. Computing semantic similarity in a group of consecutive words based on the word vectors is defined as *contextual semantic similarity*, and it could be useful in predicting naturalistic discourse reading and listening. This study proposes a novel method to compute contextual semantic similarity based on word vectors trained by discriminative learning. Specifically, we look at the preceding n words for the target word. The vectors representing the preceding words and the current word constitute a new matrix. Value in each cell of this matrix is derived from discriminative-learning word embebddings. Each cell in this matrix represents the semantic activation between two words. After that, we add up the absolute values in this lower triangular matrix under the diagonal and obtain a value. This value actually represents the contextual semantic similarity for this target and its context. It turns out that the contextual semantic similarity proposed in this study is a good predictor for the data on eye-movements, EEG signals of naturalistic discourse reading and listening. Overall, this study creates a novel semantic similarity computational model based on interpretable word vectors for word prediction in language comprehension.

Keywords: contextual semantic similarity; semantic activation; similarity computation; discriminative learning; word prediction

Modelling trial-to-trial learning in lexical decision with Linear Discriminative Learning

Maria Heitmeier, Yu-Ying Chuang and R. Harald Baayen

Throughout their lives, humans constantly update their beliefs about the world when encountering new information, which changes how they perceive and react to the world (e.g. O'Reilly and Rohrlich, 2018; Ramscar, 2016). In psycholinguistics, this is observed in priming and antipriming studies where previously observed stimuli influence the processing of subsequent stimuli. (Anti)priming can be modelled with the Rescorla-Wagner rule (Rescorla and Wagner, 1972), a mathematical account of error-driven learning (Marsolek, 2008), by assuming that the learning of the prime influences processing of the target stimulus. This implies that participants are continuously learning in priming studies, and predicts that they are also learning in each trial of other psycholinguistic experiments such as lexical decision. The present research question is therefore whether trial-to-trial learning in the mental lexicon can be detected in large-scale lexical decision experiments.

So far, computational models of trial-to-trial learning and priming in the mental lexicon used simplified semantic representations with only a few experimental items (e.g. Oppenheim et al., 2010; Lentz et al., 2021). Linear Discriminative Learning (LDL; Baayen et al., 2019) on the other hand is a model of the mental lexicon with cognitively more plausible meaning representations from distributional semantics (e.g. Mandera et al., 2017), thus providing a more fine-grained account of the mental lexicon than previous models. Moreover, it is able to model incremental learning with the Widrow-Hoff rule (Widrow and Hoff, 1960), a generalisation of the Rescorla-Wagner rule. To test our hypothesis we used data from the British Lexicon Project (BLP; Keuleers et al., 2012), a large-scale lexical decision dataset. We simulated the lexical decision experiment with LDL on a trial-by-trial basis for each subject in the BLP individually. Then, reaction times for both words and nonwords were predicted with Generalised Additive Models (GAMs; Wood, 2011), using measures — such as a stimulus' semantic density, or a learned measure of how much its sublexical cues support a word response ("yes-activation") — derived from the LDL simulation as predictors. After having developed the best models with the data of subjects 1 and 2 only, we tested the models on all other subjects. This procedure ensured the generalisabilty of our results. In order to answer our research question, we ran two simulations for each subject: one with learning updates between trials and one without. We extracted measures from both simulations, and used them as input to two GAMs. Then we compared the model fit of both models with Akaike's Information Criterion (AIC).

We found that the learning-based models showed better model fit than the non-learning models for 85% of the subjects for words (mean/SD AIC difference: 35.2/40.8), and for 94% for nonwords (55.7/45.8). A closer inspection revealed that the "yes-activation" predictor — which was only available for the learning models — to some extent contributed to better model fits of the learning GAMs compared to the non-learning ones, especially for nonwords. When "yes-activation" was removed from the learning GAMs, for words 82% of learning models were better than non-learning ones (mean/SD AIC difference: 35.2/40.7), and for nonwords 60% (3.0/24.7). This suggests that trial-to-trial learning in the mental lexicon tends to affect words more strongly, while processing of nonwords is influenced more by purely form-based sublexical learning ("yes-activation"). Furthermore, our extracted measures provide insights into processing in the mental lexicon. For example, both higher semantic density and sublexical support for a word response speed up reactions for words and slow them down for nonwords.

We conclude from these results that trial-to-trial learning in the mental lexicon can indeed be detected in large-scale psycholinguistic experiments such as the BLP. The present study therefore extends previous work applying error-driven learning to modelling trial-to-trial effects by modelling within-experiment learning with a much more detailed model of the mental lexicon that integrates distributional semantics with statistical learning.

References

Baayen, R. H., Chuang, Y.-Y., Shafaei-Bajestan, E., and Blevins, J. P. (2019). The discriminative lexicon: A unified computational model for the lexicon and lexical processing in comprehension and production grounded not in (de)composition but in linear discriminative learning. *Complexity*, 2019:1–39.

Keuleers, E., Lacey, P., Rastle, K., and Brysbaert, M. (2012). The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior research methods*, 44(1):287–304.

Lentz, T., Nixon, J. S., and van Rij, J. (2021). Temporal response modelling uncovers electrophysiological correlates of trial-by-trial error-driven learning. *PsyArXiv: 10.31234/osf.io/dg5mw*.

Mandera, P., Keuleers, E., and Brysbaert, M. (2017). Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language*, 92:57–78.

Marsolek, C. J. (2008). What antipriming reveals about priming. *Trends in Cognitive Sciences*, 12(5):176–181.

Oppenheim, G. M., Dell, G. S., and Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition*, 114(2):227–252.

O'Reilly, R. and Rohrlich, J. (2018). Deep predictive learning in vision. In *2018 Conference on Cognitive Computational Neuroscience*, pages 2018–1242.

Ramscar, M. (2016). Learning and the replicability of priming effects. *Current Opinion in Psychology*, 12:80–84.

Rescorla, R. A. and Wagner, A. R. (1972). *Classical conditioning II: current research and theory*, chapter A theory of Pavlovian conitioning: variations in the effectiveness of reinforcement and nonreinforcement, pages 64–99. Appleton-Century-Crofts, New York.

Widrow, B. and Hoff, M. (1960). Adaptive switching circuits. *1960 WESCON Convention Record Part IV*.

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)*, 73(1):3–36.

When do we learn from absent cues?

Jessie S. Nixon, Sanne Poelstra and Jacolien van Rij
jessie.nixon@uni-tuebingen.de, sannepoelstra@home.nl, j.c.van.rij@rug.nl

The Rescorla-Wagner equations (Rescorla & Wagner, 1972) have had a considerable impact on psychology, due to their success in predicting a large range of observed learning phenomena. Although originally designed to model animal learning, the Rescorla-Wagner equations have also been used to model human learning, including category learning (Gluck & Bower, 1988) and language (Baayen, Milin, Đurđević, Hendrix, & Marelli, 2011; Hoppe, van Rij, Hendriks, & Ramscar, 2020; Ramscar, Yarlett, Dye, Denny, & Thorpe, 2010) including speech acquisition (Lentz, Nixon, & van Rij, 2022; Nixon, 2018, 2020; Nixon & Tomaschek, 2020, 2021).

Yet there have also been a number of reports of learning behaviour that the Rescorla-Wagner (RW) model fails to predict, thus inspiring proposals for modifications to the model. One such study is that of Van Hamme and Wasserman (1994, VHW). In a causal rating task, participants were asked to rate the likelihood of three different foods (*cues*) as causes of an allergic reaction. On each trial, participants were told which two foods were eaten and whether the allergy occurred or not. They were asked to rate *all three* foods, including food that was not eaten on that day (trial) - the 'absent cue'. VHW found that ratings changed not only to the 'present' foods, but also to the absent food. Based on these results, they proposed an update to the RW model to include learning from absent cues. Poelstra, Nixon, and van Rij (2021) presented simulations of the VHW experiment using both the RW and the VHW version of the model. The simulations highlighted some ambiguities in the VHW experiment design and possibilities for teasing apart the models.

The present study consisted of two online experiments, a replication and an extension of VHW. The extension included a Test phase at the end, testing whether weights to Block 1 cues (e.g. 'bran') decremented over later blocks when they were absent, as predicted by VHW (Figure 1, right panel), or remained stable as predicted by RW (Figure 1, left panel). In addition, new cues were also presented at Test (purple arrow in Figure 1) to see whether the new cue was rated higher than old cues, as predicted by VHW, or lower as predicted by RW. No significant difference was found between ratings in Block 1 and test. This null result is predicted by RW, but not VHW. In addition, the new cues were rated higher than the old cues in at least some cases. Overall, the results showed support for the RW, in which learning is based on present cues, and found little support for learning from absent cues.

The results of the present study raise the intriguing question of why VHW seemed to find evidence learning from absent cues, whereas the present study did not. We propose that this difference results from specific aspects of the experiment design: the kind of incremental, implicit learning that underlies language acquisition and learning 'about the world' seems to be best modelled with present cues as input; learning from 'absent cues' can also occur under certain conditions, but seems to require recruitment of higher-level cognitive processes.

References

Baayen, R. H., Milin, P., Đurđević, D. F., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological review*, *118*(3), 438.

Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General*, *117*(3), 227.

Hoppe, D. B., van Rij, J., Hendriks, P., & Ramscar, M. (2020). Order matters! influences of linear order on linguistic category learning. *Cognitive Science*, *44*(11), e12910.

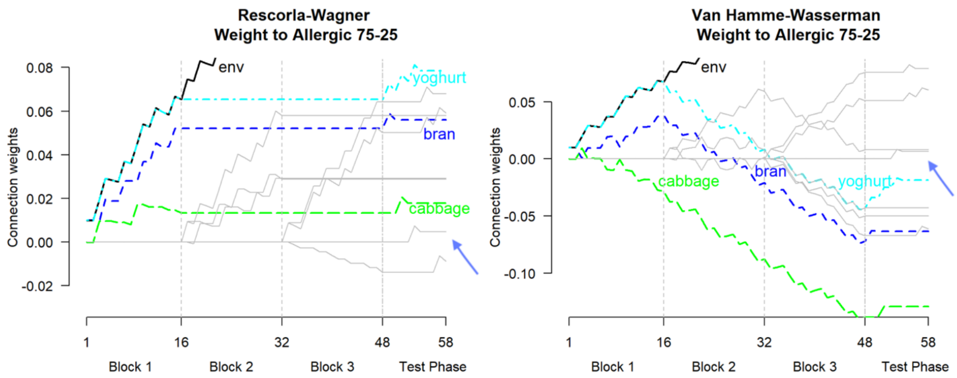


Figure 1: Simulations of the predicted response to 'allergic' using the original Rescorla-Wagner equations (**Left panel**) and the adaptation proposed by Van Hamme & Wasserman (**Right panel**). We test the predictions with a new cue after training (blue arrow). (Note that in order to maximise space in the figure, the scales are shifted relative to the vertical axis in these plots – zero is higher in VHW).

Lentz, T. O., Nixon, J. S., & van Rij, J. (2022, April). Temporal response modelling uncovers electrophysiological correlates of trial-by-trial error-driven learning. *Language, Cognition & Neuroscience*. Retrieved from https://www.psyarxiv.com/dg5mw doi: 10.31234/osf.io/dg5mw

Nixon, J. S. (2018, September). Effective acoustic cue learning is not just statistical, it is discriminative. In *Interspeech 2018 – 19th Annual Conference of the International Speech Communication Association* (pp. 1447 – 1451). Hyderabad, India.

Nixon, J. S. (2020). Of mice and men: speech acquisition as discriminative learning from prediction error, not just statistical tracking. *Cognition*, *197*, 104081.

Nixon, J. S., & Tomaschek, F. (2020). Learning from the acoustic signal:Error-driven learning of low-level acoustics discriminates vowel and consonant pairs. In S. Denison, M. Mack, Y. Xu, & B. C. Armstrong (Eds.), *Proceedings of CogSci2020, the 42nd Annual Meeting of the Cognitive Science Society* (pp. 585–591). Toronto: Cognitive Science Society.

Nixon, J. S., & Tomaschek, F. (2021). Prediction and error in early infant speech perception: A speech acquisition model. *Cognition*, *212*, 104697.

Poelstra, S., Nixon, J. S., & van Rij, J. (2021). Does learning occur in the absence of cues? In J. S. Nixon, E. Shafaei-Bajestan, & R. H. Baayen (Eds.), *Proceedings of EDLL 2021, the first International Conference on Error-driven Learning in Language* (pp. 16–17). Tübingen. Retrieved from https://quantling.org/EDLL2021/docs/EDLL2021_timetable_book_of_abstracts.pdf#page=16

Ramscar, M., Yarlett, D., Dye, M., Denny, K., & Thorpe, K. (2010). The effects of feature-label-order and their implications for symbolic learning. *Cognitive Science*, *34*(6), 909–957.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning ii: Current research and theory* (Vol. 2, pp. 64–99). New-York: Appleton-Century-Crofts.

Van Hamme, L. J., & Wasserman, E. A. (1994). Cue competition in causality judgments: The role of nonpresentation of compound stimulus elements. *Learning and motivation*, *25*(2), 127–151.

The phonetic realization of German interfixes is co-determined by semantics

Motoki Saito, R. Harald Baayen

University of Tübingen
motoki.saito@uni-tuebingen.de
harald.baayen@uni-tuebingen.de

For a long time, semantics was the most elusive aspect of language. However, developments in distributional semantics (e.g., Amenta, Crepaldi, & Marelli, 2020; Landauer & Dumais, 1997; Marelli, Amenta, & Crepaldi, 2015) have made it possible to quantify semantic similarity and investigate the consequences of semantic similarity of lexical processing and phonetic realization. A measure based on the Discriminative Lexicon Model (DLM: Baayen, Chuang, Shafaei-Bajestan, & Blevins, 2019) quantifies the amount of support that a word, or a word's sublexical units, receive from that word's meaning. Gahl and Baayen (2022) reported that the acoustic duration of English homophones is predictable to a considerable extent from the amount of semantic support that a homophone receives. Likewise, Saito, Tomaschek, and Baayen (to appear) observed, using electromagnetic articulography, that greater semantic support for the German /a/ predicted enhanced articulation of this vowel.

The present study follows up on the work by (Kuperman, Pluymaekers, Ernestus, & Baayen, 2007) on Dutch interfixes. They observed that an interfix had a longer acoustic duration when it had a greater probability in a compound's left constituent family. Kuperman et al. (2007) hypothesized that in general, a greater paradigmatic probability affords phonetic enhancement. However, it is possible that part of this enhancement is due to the semantic support that an interfix receives from the meaning of the compound in which it occurs.

We addressed this possibility for the interfixes of German. We first collected all the German compound words from the CELEX database (Baayen, Piepenbrock, & Gulikers, 1995) and calculated the values of paradigmatic probability and semantic support to the interfix. The latter measure was obtained with the algorithm for mapping from meaning to form in the discriminative lexicon model, using pre-trained embeddings based on word2vec (Müller, 2015). More specifically, we defined the semantic support for the interfix as the average support for those triphones that contain the interfix or lead into the interfix (e.g., [*nʃə*] and [*ʃən*] in *Menschenliebe* [mɛnʃənliːbə]).

Semantic support to interfix and paradigmatic probability were weakly correlated ($\tau_b = 0.073, p < 0.0001$), which opens up the possibility that semantic support to the interfix may explain variance that is not captured by paradigmatic probability. Using the Generalized Additive Mixed-effect Model (GAMM), we estimated the effects of paradigmatic probability and semantic support to interfixes on the acoustic duration of the interfix, as well as on the tongue tip positions during the articulation of the interfix. The model for duration indicated that interfix duration was longer not only for larger values of paradigmatic probability, but also for larger values of semantic support. The tongue tip model revealed similar results: a higher paradigmatic probability, and likewise a greater semantic support predicted a higher position of the tongue tip. Since the dataset contained only interfixes with alveolar consonants (i.e., /n/ and /s/), this result indicates clearer articulation for greater paradigmatic probability and greater semantic support. The relevance of semantic support as a new measure is apparent from the observation that it provides a better model fit for both of interfix duration and tongue tip positions (ΔAIC $= 3.570, \Delta$AIC $= 58.839$), when replacing paradigmatic probability as predictor.

From the general perspective of error-driven learning, it is remarkable that a measure based on a linear mapping from word embeddings to numeric vectors representing words' forms offers excellent predictivity for the phonetic detail with which words are articulated.

References

Amenta, S., Crepaldi, D., & Marelli, M. (2020). Consistency measures individuate dissociating semantic modulations in priming paradigms: A new look on semantics in the processing of (complex) words. *Quarterly Journal of Experimental Psychology*, *73*(10), 1546–1563. doi: 10.1177/1747021820927663

Baayen, R. H., Chuang, Y.-Y., Shafaei-Bajestan, E., & Blevins, J. P. (2019). The Discriminative Lexicon: A Unified Computational Model for the Lexicon and Lexical Processing in Comprehension and Production Grounded Not in (De)Composition but in Linear Discriminative Learning. *Complexity*, 1–39. doi: 10.1155/2019/4895891

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database (release 2). *Published by the Linguistic Data Consortium, University of Pennsylvania*.

Gahl, S., & Baayen, R. H. (2022). Time and thyme again: Connecting spoken word duration to models of the mental lexicon. *OSF*, 1–41. Retrieved from `https://osf.io/2bd3r/`

Kuperman, V., Pluymaekers, M., Ernestus, M., & Baayen, H. (2007). Morphological predictability and acoustic duration of interfixes in Dutch compounds. *The Journal of the Acoustical Society of America*, *121*(4), 2261–2271. doi: 10.1121/1.2537393

Landauer, T. K., & Dumais, S. T. (1997). A Solution to Plato's Problem: The Latent Semantic Analysis Theory of Acquisition, Induction, and Representation of Knowledge. *Psychological Review*, *104*(2), 211–240.

Marelli, M., Amenta, S., & Crepaldi, D. (2015). Semantic transparency in free stems: The effect of Orthography-Semantics Consistency on word recognition. *Quarterly Journal of Experimental Psychology*, *68*(8), 1571–1583. doi: 10.1080/17470218.2014.959709

Müller, A. (2015). *Analyse von Wort-Vektoren deutscher Textkorpora* (Bachelor's Thesis, Technische Universität Berlin). Retrieved from `https://devmount.github.io/GermanWordEmbeddings`

Saito, M., Tomaschek, F., & Baayen, R. H. (to appear). Articulatory effects of frequency modulated by inflectional meanings.. doi: 10.31234/osf.io/8exk7

Structure prediction occurs when it is needed: Evidence from Visual-world Structural priming in Mandarin and Dutch comprehension

Xuemei Chen[1], Suiping Wang[1], and Robert J. Hartsuiker[2]
[1] Philosophy and Social Science Laboratory of Reading and Development in Children and Adolescents (South China Normal University), Ministry of Education, 510631 Guangzhou, China
[2] Department of Experimental Psychology, Ghent University, Belgium

Arai et al. (2007) showed that structural priming in the comprehension of English dative sentences only occurred when the verb was repeated between prime and target, suggesting a lexically-dependent mechanism of structure prediction. However, structural priming studies in production of both English and Dutch have demonstrated stronger priming effects for unexpected sentence structures (e.g., a verb biased towards double object (DO) structure is followed by an unexpected prepositional object (PO) structure). The latter finding of *inverse preference* priming is consistent with error-based implicit learning accounts that suggest structural priming in both production and comprehension is based on learning from prediction errors (Chang et al., 2006). Therefore, it is unclear which mechanism holds for structure prediction in comprehension (lexically-dependent vs. abstract), and whether such a mechanism is shared with production. Here we tested the structural priming in visual-world comprehension experiments in Dutch and Mandarin Chinese. Both languages are biased towards the PO structure while Dutch is a Germanic language like English and Mandarin is a Sino-Tibetan language.

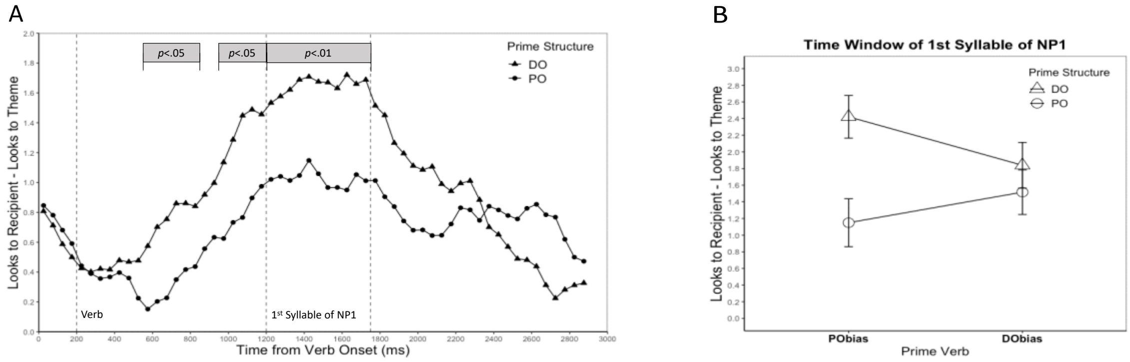
In the study of Mandarin comprehension, Experiment 1 was a norming study (N=367) that measured the biases (DO vs. PO) of 48 Mandarin Chinese dative verbs. Experiment 2 (N=72) crossed verb bias (DO-bias or PO-bias) and structure (DO or PO) of prime sentences in a visual-world paradigm, to examine whether Mandarin comprehenders show an inverse preference effect. The priming effect is expressed as the proportion of looks to the predicted referent (i.e., the recipient after a DO-prime, the theme after a PO-prime), for two critical time windows during target sentence processing: the verb and the first syllable of the first post-verbal noun (which was identical in theme and recipient). There was priming in both time windows, even though the verb differed between prime and target (Figure 1A). Importantly, there was an inverse preference effect (i.e., stronger priming after a DO prime with a PO-biased verb than with a DO-biased verb) in the second time window (Figure 1B). These results provide evidence for an error-based structure prediction system in comprehension.

Similar to the Mandarin study, we constructed a pretest of verb bias and then tested both structural priming and inverse preference priming in the study of Dutch comprehension. However, we did not find the inverse preference priming effect and the abstract structural priming regardless of the speech rate of the target sentences in Experiments 1 and 2 (Figure 2A, 2B). In contrast, structural priming only occurred when the verb was repeated (Experiment 3, N=36, Figure 2C) or when the verb was different but the target structure was relatively unpredictable (Experiment 4, N=36, Figure 2D). We interpret the inconsistent findings across languages in terms of an effortful process of structure prediction in comprehension (Pickering & Gambi, 2018): Prediction occurs when it is needed (e.g., in Mandarin and in Dutch when the target structure is unpredictable), but not when it is not necessary (e.g., when the target structure is predictable in Dutch).

Key words: Implicit learning; verb bias; structural priming; comprehension; prediction

Figure 1

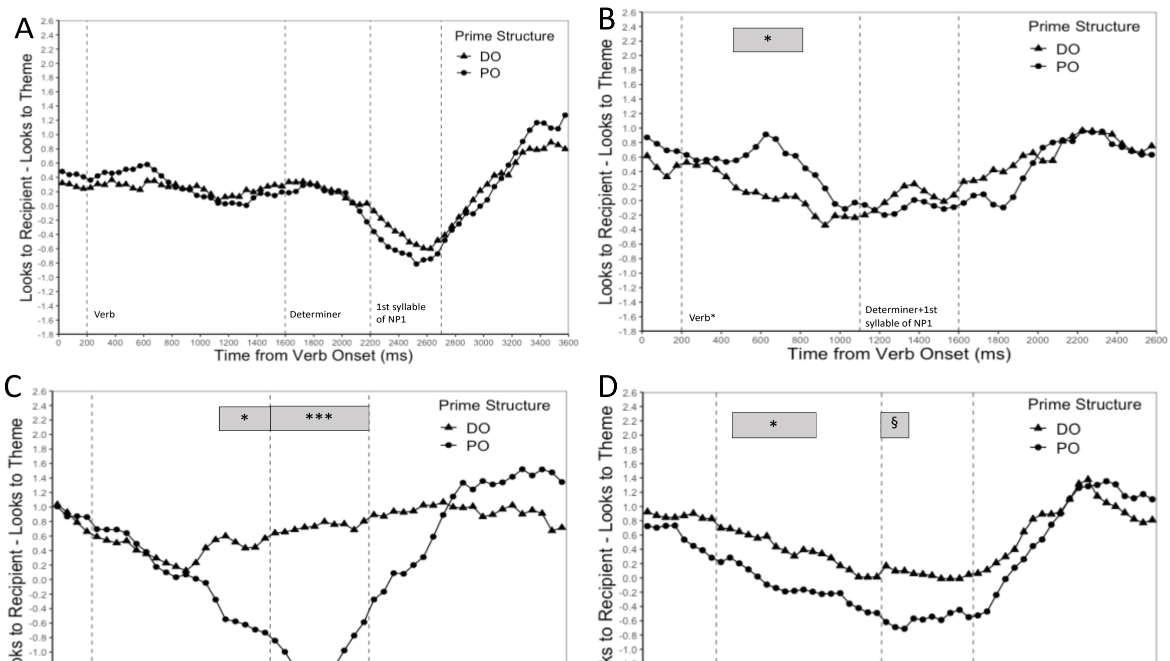
Difference in proportion of looks to recipient and theme for each time bin (50ms) from onset of target verb in Experiment 2 of Mandarin study



Note. The time window of verb is from 200ms to 1200ms and the time window of the first syllable of the first noun phrase is from 1200ms to 1750ms. Plot A on the left indicates the difference in the proportions of looks to recipient (predicting DO structure) and to theme (predicting PO structure) after prime sentences with different structure (DO vs. PO); the grey rectangles within the time windows indicate the clusters where the main effect of structure was significant (by-subject analysis). Plot B on the right indicates the interaction between structure and verb bias in the time window of the first syllable of the first noun phrase (p<.05). The error bars reflect standard errors from a by-participant analysis.

Figure 2

Difference in proportion of looks to recipient and theme for each time bin (50ms) from onset of target verb in the Experiments of Dutch study



Note. Four plots indicate the difference in the proportions of looks to the recipient (predicting DO structure) and to the theme (predicting PO structure) after prime sentences with different structure (DO vs. PO). Plot A on the top left indicates priming in Experiment 1 with different verb and slow speech of target sentences. The time window of the verb is from 200ms to 1600ms, the time window of the determiner is from 1600ms to 2200ms, and the time window of the first syllable of the first noun phrase is from 2200ms to 2700ms. Plot B, C, and D indicate priming with normal speech of target sentences in Experiments 2, 3, and 4. The time window of the verb is from 200ms to 1100ms, the time window of the determiner and the first syllable of the first noun phrase is from 1100ms to 1600ms. Plot B on the top right indicates priming in Experiment 2 with different verb and predictable target structure. Plot C on the bottom left indicates priming in Experiment 3 with same verb and predictable target structure. Plot D on the bottom right indicates priming in Experiment 4 with different verb but unpredictable target structure. The marker "*" in the label of the verb time window indicates the main effect of prime structure. The grey rectangle within the time window indicate the cluster where the main effect of prime structure was significant (by-subject analysis). §p <.1, *p < .05, **p < .01, ***p < .001.

Does structural priming lead to learning? Evidence from classroom learners of Mandarin

Yanxin (Alice) Zhu, Theres Grüter
University of Hawai'i

The effects of structural priming have been viewed as a reflection of error-based implicit learning (Chang et al., 2006). In this tradition, evidence for "learning" is typically seen in increased production of the primed construction(s), i.e., adaptation of the output towards input distributions (Jaeger & Snider, 2013). By contrast, evidence for "learning" in second language acquisition (SLA) often comes from more nativelike acceptability judgments (e.g., Tachihara & Goldberg, 2020). As structural priming is gaining increased attention as a potential learning mechanism in SLA (Jackson, 2018), there is need to connect these different views of "learning." As a first step towards bridging this gap, this study investigates whether structural priming could facilitate L2 learning of the Mandarin dative alternation (see Table 1) as reflected in (i) increased production of acceptable verb-dative combinations and decreased production of unacceptable ones, and (ii) increased acceptability ratings for acceptable and decreased ratings for unacceptable combinations, following a priming task in which only acceptable constructions were primed.

Methods Participants were 25 native speakers (L1) and 41 classroom learners (CL) of Mandarin (LexTALE_CH score (/90, Chan & Chang, 2018: L1=78.6 (69-86), CL=57.4 (43-75)). The experiment implemented a pretest-treatment-posttest design (Fig1). During treatment (priming; B2, Fig1), participants were exposed to acceptable verb-dative combinations only. We compared participants' production (B1 vs. B3, Fig1) and judgments (A vs. C, Fig1) before and after priming to assess learning.

Results Separate Logistic Mixed-Effects Regression models were conducted for productions with each verb type, with GO vs. non-GO as the dependent variable. No changes were observed with TELL verbs since (acceptable) DO productions were at ceiling from the start. The CL group produced significantly more (acceptable) GO constructions with MAKE verbs post- vs pre-priming (Fig2, b=8.4, p<.001), while the L1 group increased MAKE-GO productions only numerically (b=4.3, p=.16). With GIVE verbs, for which the priming phase contained an equal number of DO and GO constructions, GO productions increased in the CL group(b=1.6, p<.001); this differential increase in GO vs. DO productions aligns with error-based learning accounts, which predict greater change for less expected constructions. We assume GO to be less expected than DO with GIVE verbs for CL learners since introductory Chinese textbooks use GIVE verbs almost exclusively with DO (for reasons that remain to be determined).

Fig3 presents ratings for the four acceptable and two unacceptable (see Table 1) verb-dative pairings in the AJTs before and after priming. LMER models indicate ratings of acceptable pairings increased in both groups (L1: b=0.18, p<.001; CL: b=0.45, p<.001), yet more so in the CL group (interaction: b=0.27, p<.01). These results are consistent with positive evidence supporting entrenchment, and the CL group's less entrenched representations being more susceptible to change following positive evidence. Unexpectedly, ratings for unacceptable pairings also increased (L1: b=0.38, p<.001; CL: b=0.18, p=.01), with no significant difference between groups. We thus see no evidence of indirect negative evidence or statistical preemption (Goldberg, 2006) leading to decreased judgments of unacceptable pairings.

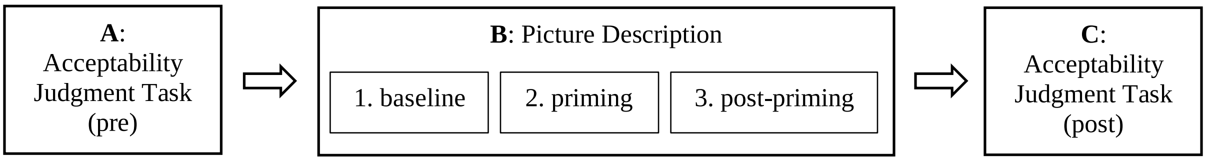
Conclusions These findings can be taken as evidence for learning in both senses in that structural priming led to increased (and more nativelike) *production* of acceptable pairings (GO with MAKE and GIVE verbs) as well as increased *ratings* of acceptable pairings overall. Taking these changes as reflecting changes in learners' linguistic representations also aligns with the notion that language representations are not stable but rather continuously updated based on recent experience (Jaeger & Snider, 2013). On the other hand, these results do not present evidence for priming leading to learning what is unacceptable (preemption).

References
Chan, IL & Chang, CB (2018). LEXTALE_CH: A quick, character-based proficiency test for Mandarin Chinese. In *Proceedings of BUCLD 42*. Cascadilla.
Chang, F, Dell, GS & Bock, K (2006). Becoming syntactic. *Psychological Review, 113*.
Goldberg, AE (2006). *Constructions at work: the nature of generalization in language.* OUP.
Jackson, CN (2018). Second language structural priming. *Second Language Research, 34.*
Jaeger, TF & Snider, NE (2013). Alignment as a consequence of expectation adaptation. *Cognition, 127*.
Liu, F (2006). Dative constructions in Chinese. *Language and Linguistics, 7*.
Tachihara, K, & Goldberg, A (2020). Reduced competition effects and noisier representations in a second language. *Language Learning, 70*.

Table 1. Double-object (DO) and GEI-object (GO) dative constructions in Mandarin illustrated with verbs from the three semantic verb classes used in the experiment; acceptable verb-construction pairings in blue, unacceptable ones in red (Liu, 2006).

	Double-object (DO) dative	***GEI-object (GO) dative***
MAKE	*Mali **zuo** le Dawei yi ge dangao. Mary **make** ASP David a CL cake 'Mary **made** David a cake.'	Mali **zuo** le yi ge dangao gei Dawei. Mary **make** ASP a CL cake GEI David 'Mary **made** a cake for David.'
TELL	Mali **gaosu** le Dawei yi ge mimi. Mary **tell** ASP David a CL secret 'Mary **told** David a secret.'	*Mali **gaosu** le yi ge mimi gei Dawei. Mary **tell** ASP a CL secret GEI David 'Mary **told** a secret to David.'
GIVE	Mali **song** le Dawei yi ge dangao. Mary **give** ASP David a CL cake 'Mary **gave** David a cake.'	Mali **song** le yi ge dangao gei Dawei. Mary **give** ASP a CL cake GEI David 'Mary gave a cake to David.'

Figure 1. *Overview of Procedure*



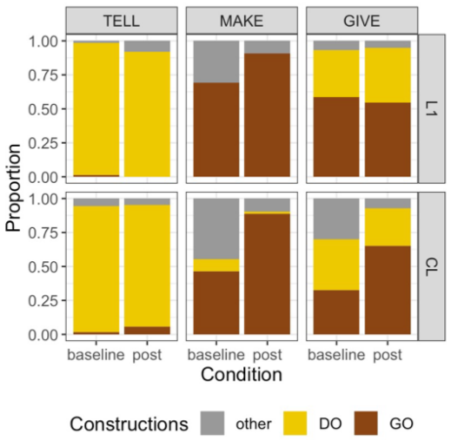


Figure 2. Proportion DO, GO and other constructions produced by the two groups pre- and post-priming

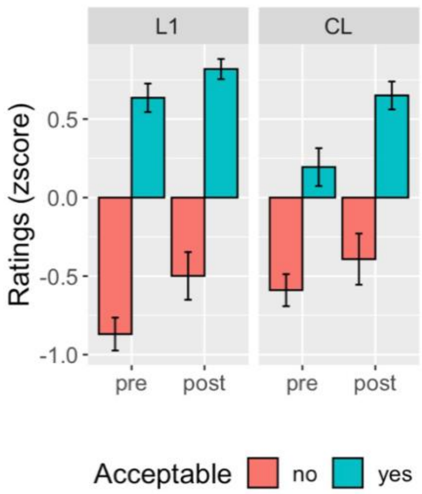


Figure 3. Ratings for acceptable (blue) and unacceptable (red) verb-construction pairings by group pre- and post-priming

The effect of verb surprisal on the acquisition of new syntactic structures

Giulia Bovolenta and Emma Marsden (University of York)

Error-based learning models of language acquisition such as the Dual Path model (Chang et al., 2006) can reproduce phenomena from L1 syntax acquisition as well as inverse frequency priming effects in adults (i.e., the fact that encountering a syntactic structure with verbs not normally used with it leads to enhanced structural priming). Such findings suggest that L1 acquisition and processing may be driven by prediction error: as we encounter unexpected input, we update our representations to match that input. Inverse frequency priming effects can also be observed in L2 speakers (Montero-Melis & Jaeger, 2020), further suggesting that error-based learning mechanisms may still be active during L2 acquisition. However, there is so far little evidence of a direct role of prediction error in second language acquisition. Can prediction error—specifically, encountering verbs in unexpected syntactic contexts—drive the acquisition of new syntactic structures in adults?

Design and Methods. We carried out two pre-registered experiments (Exp 1: N = 86; Exp 2: N = 84) with native English speakers. Over three days, participants were exposed to an artificial language (Table 1) with two syntactic structures which differed in word order, verb inflection and preposition use: active (e.g., *Lu meeb shen**at** lu prag,* 'The girl calls the boy') and passive (e.g., *Lu prag shen**es ka** lu meeb,* 'The boy is called by the girl'). Participants were trained using a cross-situational learning paradigm (Rebuschat et al., 2021) with 288 photographs depicting transitive actions. They heard individual sentences while two pictures (a target and a distractor) appeared on screen side by side; they then had to select the picture that corresponded to the sentence they had just heard (the target). They received no feedback. Responses were initially at chance level, but participants gradually learned the language as they encountered more sentences. Training on Day 1 established expectations for specific verbs to only occur with one of the two structures. The strength of this expectation varied between experiments. In Experiment 1, 8 out of 12 verbs on Day 1 were structure-specific, while others could occur with both structures; in Experiment 2, participants were only trained on the 8 structure-specific verbs. In both experiments, on Day 2, the Surprisal group encountered sentences which placed structure-specific verbs within an unexpected structure, while the Control group received input that matched the Day 1 pattern.

Outcome measures. Participants' structural comprehension was tested during training on Day 2 (to test for immediate priming) using a modified version of the training paradigm. We manipulated the target and distractor pictures so that they only differed by agent and patient roles (e.g., one picture depicted a boy calling a girl, the other picture a girl calling a boy), forcing participants to interpret the syntactic structure to select the correct picture. On Day 3, participants were tested again on structural comprehension and on a grammaticality judgment task (where incorrect items had the active inflection with *ka* or the passive inflection without *ka*). Response accuracy in structural comprehension was entered in logistic mixed-effects models with group and structure as fixed effects, while *d'* scores for grammaticality judgments were entered in linear models with group and structure as predictors.

Results. On Day 2, we observed no evidence of enhanced immediate priming from surprising sentences in either experiment. But on Day 3, results showed an effect of surprisal on structural comprehension and grammaticality judgment accuracy, which appeared to be driven by the strength of the initial expectation for verbs to be structure-specific: in Experiment 1 (that generated weaker expectations) we observed trends, which became significant effects in Experiment 2 (that generated stronger expectations). Specifically, participants in the Surprisal group showed better structural comprehension and higher accuracy on the grammaticality judgment task for one structure (passive), suggesting that they had developed stronger structural representations from encountering novel syntactic structures with unexpected verbs (Figures 1 & 2). These findings suggest that the acquisition of new structures by adults may be driven by prediction error, providing evidence of error-based learning in early L2 acquisition.

Table 1. Summary of the artificial language used. Individual word-meaning assignment within word categories was randomised by participant.

Category	Words	Meanings
Noun	glim, blom, prag, meeb	*man, woman, boy, girl*
Verb stem	flug-, loom-, gram-, pod-, zal-, shen-, norg-, klig-, jeel-, lemb-, gond-, vang-	*call, chase, greet, interview, pay, photograph, scare, threaten, dismiss, serve, kick, tease*
Determiner	lu	*the*
Preposition	ka	*by*

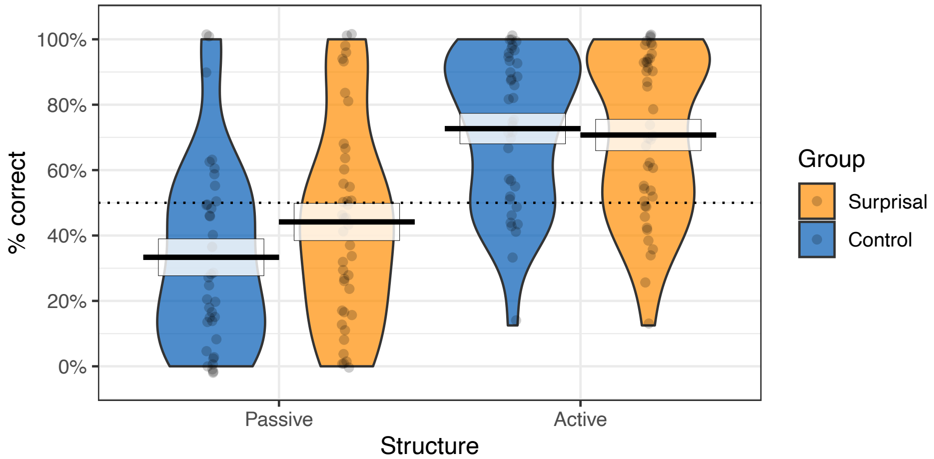


Figure 1. Mean accuracy, Day 3 structure comprehension test (Experiment 2). Shaded rectangles: 95% CI; horizontal dotted line: chance performance (50%).

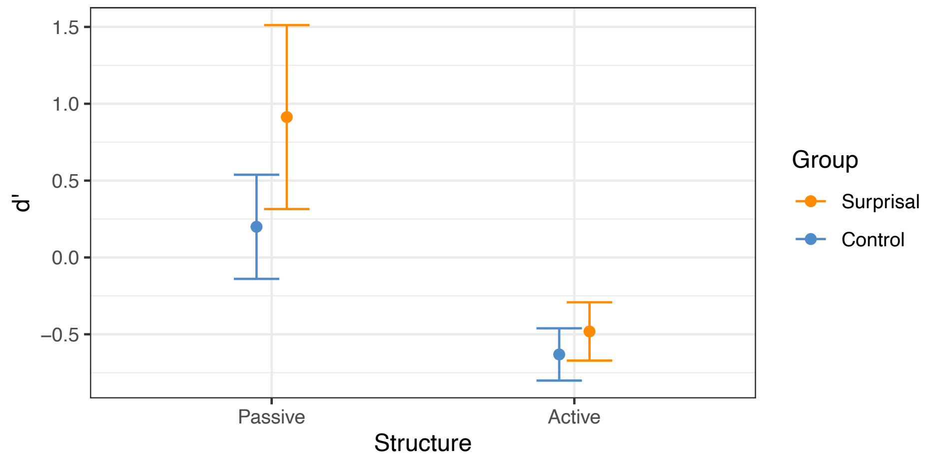


Figure 2. Mean *d'* scores, grammaticality judgment task (Experiment 2). Error bars: 95% CI.

References

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, *113*(2), 234–272.

Montero-Melis, G., & Jaeger, F. T. (2020). Changing expectations mediate adaptation in L2 production. *Bilingualism: Language and Cognition*, *23*(3), 602–617.

Rebuschat, P., Monaghan, P., & Schoetensack, C. (2021). Learning vocabulary and grammar from cross-situational statistics. *Cognition*, *206*, 104475.

The babe with the predictive power: work in progress examining the role of prediction in early word encoding

Judit Fazekas[a], Yamil Vidal[b], Julian Pine[c] & Perrine Brusini[c]

[a]School of Psychological Sciences, University of Manchester, Ellen Wilkinson Building, Manchester M13 9PL, UK
[b] Donders Institute for Brain, Cognition and Behaviour, Radboud University, Kapittelweg 29, 6525 EN, Nijmegen, Netherlands.
[c]Department of Psychological Sciences, University of Liverpool, Eleanor Rathbone Building, Liverpool L69 7ZA, UK

Introduction: Error-based theories of language acquisition posit that predictions are a key part of language processing throughout the human lifespan. They suggest that adults and children are constantly anticipating upcoming input, assess these predictions and that they use the potential discrepancies to update their linguistic knowledge. In order to describe the earliest stages of language acquisition, such theories hold that infants start making predictions soon after birth. However, linguistic predictions are challenging to target experimentally, and existing studies have typically focused on linguistic prediction in older age groups, using methods that rely on participants' advanced linguistic knowledge. As a result, there is currently limited evidence that prediction is a viable learning mechanism in infancy and thus that error-based language acquisition can operate throughout the whole lifespan. Focusing on these initial stages of predictive language development, the current study targets the role of prediction in early word encoding. In particular, we will assess the acquisition of word-forming sound sequences and 9-month-olds' capacity to predict specific upcoming sounds when listening to familiar words.

Methods: We have adapted an adult EEG study focusing on syllabic prediction (Vidal et al., 2019) for an infant population. Our study starts with a learning phase, in which 39 nine-month-old monolingual English-speaking infants hear two trisyllabic pseudowords. These words are then used as standard stimuli in an oddball-phase with four new words. Two of these deviant words only share their first syllable with a familiar word while the other two share their first two syllables.

We will measure infants' mismatch-response (MMR), an early-onset EEG event related potential that tends to occur when the expected event and the one being processed do not align (e.g., Friston, 2005). The MMR is associated with failed predictions (Wacongne et al., 2011), which play a crucial role in prediction-based language acquisition theories. Thus, the MMR could prove to be instrumental in both targeting linguistic predictions themselves, and in examining their potential role in language learning (e.g., Partanen et al., 2013). We will assess whether 9-month-olds' MMR differs between standard and deviant words, in order to assess whether they make phonemic-level predictions. We will also assess the MMR-difference between the two kinds of deviants. An MMR difference after one versus two shared syllables would suggest that cumulative congruent input reinforces prediction.

As infants' MMR can vary, we will also carry out a second task to localize the individual MMR responses of each participant in the form of a tone-change-detection Optimum-1 task. This task will determine the location, latency and polarity of the MMR for each infant separately, and will help ensure that the study has sufficient statistical power.

Analysis: We will use the data from the localizer task to determine the latency, location and polarity of the MMR for each participant. We will then include the average voltage for each trial in the main task in linear mixed effect models to assess the following two hypotheses:

> **Hypothesis 1: Infants can already make auditory predictions at 9 months of age.** We will assess whether prediction error, as measured by the amplitude of the MMR, is different in congruent and incongruent pseudowords.

Hypothesis 2: 9-month-old infants' prediction-error-response is sensitive to the amount of congruent evidence that occurs prior to the expectation violation. We will assess whether the amplitude of the MMR is different after expectation violation following one or two congruent syllables.

Pilot data collection is currently underway and we expect to have preliminary results by Autumn 2022.

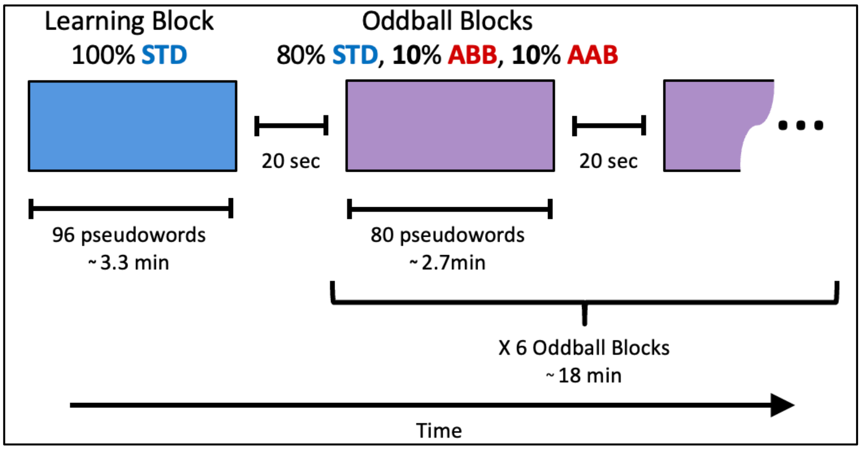


Figure 1 - General study design

Set 1 – word 1	vi:	lu:	dɑ:
Set 1 – word 2	θl	sɔ:	næ
Set 2 – word 1	vi:	sɔ:	næ
Set 2 – word 2	θl	lu:	dɑ:
Set 3 – word 1	vi:	lu:	næ
Set 3 – word 2	θl	sɔ:	dɑ:

Table 1: Word sets for the main study. An equal number of infants will hear set 1, 2 or 3 in the learning phase, while the remaining two sets will be featured as deviants in the oddball phase.

References:

Friston, K. (2005). A theory of cortical responses. Philosophical transactions of the Royal Society B: Biological Sciences, 360(1456), 815-836.

Partanen, E., Pakarinen, S., Kujala, T., & Huotilainen, M. (2013). Infants' brain responses for speech sound changes in fast multifeature MMN paradigm. Clinical Neurophysiology, 124(8), 1578-1585.

Vidal, Y., Brusini, P., Bonfieni, M., Mehler, J., & Bekinschtein, T. A. (2019). Neural signal to violations of abstract rules using speech-like stimuli. Eneuro, 6(5).

Wacongne, C., Labyt, E., van Wassenhove, V., Bekinschtein, T., Naccache, L., & Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. Proceedings of the National Academy of Sciences, 108(51), 20754-20759.

Using linear discriminative learning to model the acoustic duration of English derived words

Simon David Stein[1], Ingo Plag[1]
[1]Heinrich Heine University Düsseldorf

Recent findings in morpho-phonetic and psycholinguistic research have indicated that phonetic detail can vary by morphological structure. For example, acoustic duration seems to be affected by a segment's morphological function, affix informativeness, or decomposability (Plag et al. 2020; Tomaschek et al. 2019; Seyfarth et al. 2017; Plag et al. 2017; Ben Hedia 2019; Zuraw et al. 2020; Hay 2007, 2003). However, the mechanisms behind these effects are unclear, because traditional and current models of the morphology-phonology interaction and of speech production cannot account for these findings (e.g., Turk & Shattuck-Hufnagel 2020; Roelofs & Ferreira 2019; Levelt et al. 1999; Dell 1986; Kiparsky 1982). These models are either underspecified regarding the processing of complex words, or do not allow for post-lexical access of morphological information.

Error-driven learning offers an opportunity to explore alternative explanations for morpho-phonetic variation and the processing of complex words. One promising approach relying on error-driven learning is the Discriminative Lexicon (Baayen et al. 2019). In discriminative approaches, differences between morphological functions are expected to emerge naturally from sublexical and contextual cues. Recently, Tomaschek et al. (2019) and Schmitz et al. (2021) successfully employed a discriminative approach to model the duration of English inflectional [s] and [z]. The present paper applies the discriminative approach to derivational morphology.

Our study investigates the durational properties of 4530 English derivative tokens with the derivational functions NESS, LESS, IZE, ATION and DIS from the Audio BNC (Coleman et al. 2012). We predict the duration of these derivatives using multiple linear regression models and mixed-effects models. The crucial predictors in our models are derived from linear discriminative learning (LDL) networks (Baayen et al. 2019). These networks were trained on semantic vectors from the TASA corpus (Ivens & Koslin 1991), which had been incrementally learned through a Widrow-Hoff learning algorithm (Widrow & Hoff 1960). We construct three LDL networks which differ in the amount of morphological information shared between their vectors. This way, we can explore to which degree morphological structure needs to be encoded in the discriminative lexicon in order to explain phonetic variation.

We find, first, that measures derived from all three LDL networks are successful in predicting duration and often seem to capture similar underlying dimensions as traditional variables, such as frequency and semantic transparency. Second, a detailed inspection of the correlation strength of a word's predicted semantics with the semantics of its neighbors shows that the distributions of these correlations reflect the morphological functions NESS, LESS, IZE, ATION and DIS. These functions emerge even from a network that we do not explicitly provide with separate vectors for these functions. This implies that for modeling speech production, we do not need to explicitly encode morphological structure in our lexicon. Such structure can simply be emergent, in the sense that it is not required as part of the generating system (cf. Baayen et al. 2018). Third, the effects of several LDL measures on duration can give us new insights into mechanisms of speech production. For example, we find that words are lengthened when the semantic support of the word's predicted articulatory path is stronger.

Our results show that error-driven learning has the power to let derivational functions emerge as a by-product of a system without explicitly encoded morphological structure. In addition, approaches relying on error-driven learning can successfully model the acoustic properties of complex words. The Discriminative Lexicon, in contrast to traditional models which do not allow for a direct morphology-phonetics interaction, is well-equipped to explain this interaction by its underlying principles of learning by experience. The present study thus contributes to the growing literature that demonstrates linear discriminative learning to be a promising alternative approach to speech production which can explain phonetic variation in simplex, complex, or non-words (cf. Schmitz et al. 2021; Chuang et al. 2020; Baayen et al. 2019).

References

Baayen, R. Harald, Yu-Ying Chuang & James P. Blevins. 2018. Inflectional morphology with linear mappings. *The Mental Lexicon* 13.2: 230–268. doi: 10.1075/ml.18010.baa.

Baayen, R. Harald, Yu-Ying Chuang, Elnaz Shafaei-Bajestan & James P. Blevins. 2019. The discriminative lexicon: A unified computational model for the lexicon and lexical processing in comprehension and production grounded not in (de)composition but in linear discriminative learning. *Complexity* 2019: 1–39. doi: 10.1155/2019/4895891.

Ben Hedia, Sonia. 2019. *Gemination and degemination in English affixation: Investigating the interplay between morphology, phonology and phonetics* (Studies in Laboratory Phonology 8). Berlin: Language Science Press. doi: 10.5281/zenodo.3232849.

Chuang, Yu-Ying, Marie L. Vollmer, Elnaz Shafaei-Bajestan, Susanne Gahl, Peter Hendrix & R. Harald Baayen. 2020. The processing of pseudoword form and meaning in production and comprehension: A computational modeling approach using linear discriminative learning. *Behavior Research Methods* 53: 945–976. doi: 10.3758/s13428-020-01356-w.

Coleman, John, Ladan Baghai-Ravary, John Pybus & Sergio Grau. 2012. *Audio BNC: The audio edition of the Spoken British National Corpus.* University of Oxford. http://www.phon.ox.ac.uk/AudioBNC.

Dell, Gary S. 1986. A spreading-activation theory of retrieval in sentence production. *Psychological Review* 93.3: 283–321. doi: 10.1037/0033-295X.93.3.283.

Hay, Jennifer. 2003. *Causes and consequences of word structure.* New York, London: Routledge. doi: 10.4324/9780203495131.

Hay, Jennifer. 2007. The phonetics of *un*. In Judith Munat (ed.), *Lexical creativity, texts and contexts* (Studies in Functional and Structural Linguistics 58), 39–57. Amsterdam, Philadelphia: John Benjamins. doi: 10.1075/sfsl.58.09hay.

Ivens, Stephen H. & Bertram L. Koslin. 1991. *Demands for reading literacy require new accountability measures.* Brewster: Touchstone Applied Science Associates.

Kiparsky, Paul. 1982. Lexical morphology and phonology. In In-Seok Yang (ed.), *Linguistics in the morning calm: Selected papers from SICOL*, 3–91. Seoul: Hanshin.

Levelt, William J. M., Ardi Roelofs & Antje S. Meyer. 1999. A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22.1: 1–38. doi: 10.1017/S0140525X99001776.

Plag, Ingo, Julia Homann & Gero Kunter. 2017. Homophony and morphology: The acoustics of word-final S in English. *Journal of Linguistics* 53.1: 181–216. doi: 10.1017/S0022226715000183.

Plag, Ingo, Arne Lohmann, Sonia Ben Hedia & Julia Zimmermann. 2020. An <s> is an <s’>, or is it? Plural and genitive-plural are not homophonous. In Lívia Körtvélyessy & Pavol Štekauer (eds.), *Complex words: Advances in morphology*, 260–292. Cambridge: Cambridge University Press. doi: 10.1017/9781108780643.015.

Roelofs, Ardi & Victor S. Ferreira. 2019. The architecture of speaking. In Peter Hagoort (ed.), *Human language: From genes and brains to behavior*, 35–50. Cambridge, Massachusetts: MIT Press.

Schmitz, Dominic, Ingo Plag, Dinah Baer-Henney & Simon D. Stein. 2021. Durational differences of word-final /s/ emerge from the lexicon: Modelling morpho-phonetic effects in pseudowords with linear discriminative learning. *Frontiers in Psychology* 12: 680889. doi: 10.3389/fpsyg.2021.680889.

Seyfarth, Scott, Marc Garellek, Gwendolyn Gillingham, Farrell Ackerman & Robert Malouf. 2017. Acoustic differences in morphologically-distinct homophones. *Language, Cognition and Neuroscience* 33.1: 32–49. doi: 10.1080/23273798.2017.1359634.

Tomaschek, Fabian, Ingo Plag, Mirjam Ernestus & R. Harald Baayen. 2019. Phonetic effects of morphology and context: Modeling the duration of word-final S in English with naïve discriminative learning. *Journal of Linguistics* 57.1: 1–39. doi: 10.1017/S0022226719000203.

Turk, Alice & Stefanie Shattuck-Hufnagel. 2020. *Speech timing: Implications for theories of phonology, speech production, and speech motor control.* New York: Oxford University Press. doi: 10.1093/oso/9780198795421.001.0001.

Widrow, Bernard & Marcian E. Hoff. 1960. Adaptive switching circuits. *IRE WESCON Convention Record Part IV*: 96–104. doi: 10.21236/AD02.

Zuraw, Kie, Isabelle Lin, Meng Yang & Sharon Peperkamp. 2020. Competition between whole-word and decomposed representations of English prefixed words. *Morphology* 31: 201–237. doi: 10.1007/s11525-020-09354-6.

Learning to unlearn: The role of negative evidence in morphophonological learning

Erdin Mujezinović & Ruben van de Vijver
Heinrich-Heine University Düsseldorf

Language learning within linguistic theory is commonly conceived of as frequency-based, with only positive evidence impacting learning. Yet, many effects within language learning are hard to explain without assuming a role for negative evidence (Nixon, 2020, Ramscar et al., 2010). Within discriminative error-driven learning (Hoppe et al., 2022, Ramscar, 2021, Rescorla and Wagner, 1972) negative evidence plays an important role: The association between cues and outcomes is continuously being updated, and there is competition among cues as predictors for outcomes. This cue competition leads to increased weights of a cue-outcome association (positive evidence), if a cue correctly predicts an outcome, but also to decreased weights (negative evidence), when a cue is expected to predict an outcome, but the outcome is not present. Negative evidence then causes the *unlearning* of a cue-outcome relation. Especially within morphophonological learning, unlearning is often needed, as a previously learned phonological form does not always denote the same grammatical function (e.g. the final syllable in [plaːgə] "plague" does not discriminate among plural meanings, as it does in [taːgə] "days"). Unlearning has been proposed as a central mechanisms responsible for the the disappearance of (morphological) overregularization during language development (Ramscar, Dye, and McCauley, 2013).

We tested whether unlearning based on negative evidence impacts regular morphophonological learning in a series of artificial language learning experiments: 120 participants across four groups were taught grammatical functions (singular, plural (PL), diminutive (DM)), within two conditions: DM_Unlearning and PL_Unlearning. In DM_Unlearning, two groups were presented with singular (e.g. [be.fal]) and plural forms (e.g. [bef.li]), while in PL_Unlearning, two groups heard diminutive [bef.lit͡ɕ]) instead of plural form: All four were additionally exposed to diminutive-plural forms (e.g. [bef.li.t͡ɕi]). The critical manipulation was the presentation order, as each group within condition received a different presentation order: One group heard single diminutive/plural forms and their singulars first, and diminutive-plural forms and their singulars second. The other group was exposed to the reversed order. Learning simulations computed with `edl` (van Rij and Hoppe, 2021) on the basis of the Rescorla-Wagner learning rule predict that only the groups exposed to diminutive-plural forms first will learn the actual discriminative phonological cues (`ić` [it͡ɕ] for diminutive in DM_Unlearning, `i#` [i] for plural in PL_Unlearning) by unlearning the non-discriminative ones (`i#` [i] for diminutive, `ić` [it͡ɕ] for plural). The 120 human participants were asked to classify diminutive and plural wordforms in a 2AFC-task.

The results for DM_Unlearning are shown in Fig. 1 and for PL_Unlearning in Fig. 2. While both groups in DM_Unlearning learned the withheld diminutive-category equally well, differences between presentation orders still arose, as the plural-category instead was better learned within the group for which unlearning was available (diminutive-plural first). The results for the PL_Unlearning condition however clearly confirmed a positive effect of unlearning in learning the withheld plural-category, with clear differences in accuracy between diminutive-first and diminutive-plural first groups. Unlearning the non-discriminative `ić` [it͡ɕ] cue allowed the correct association between `i#` [i] and plural to emerge. The differences between conditions likely reflect the differences in salience between `i#` [i] and `ić` [it͡ɕ], with the latter being more salient, resulting in the former being overshadowed. In case no unlearning was available, the correct plural-category cue was learned worse in both conditions. All in all these results are in line with error-driven learning, and confirm that morphophonological learning benefits from negative evidence.

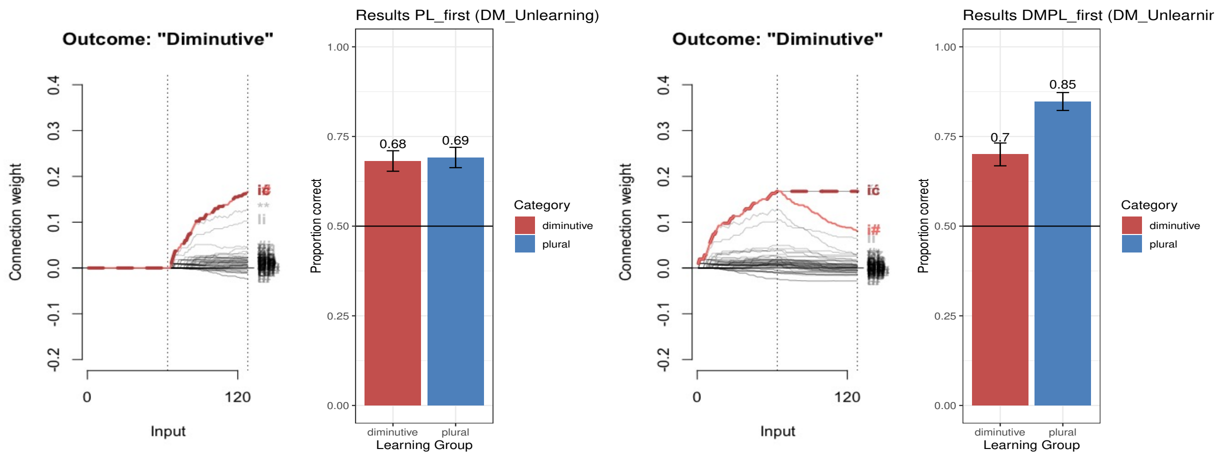


Figure 1: Results of the simulation (left most plots of the two pairs) and experimental results (right most plots) for DM_Unlearning

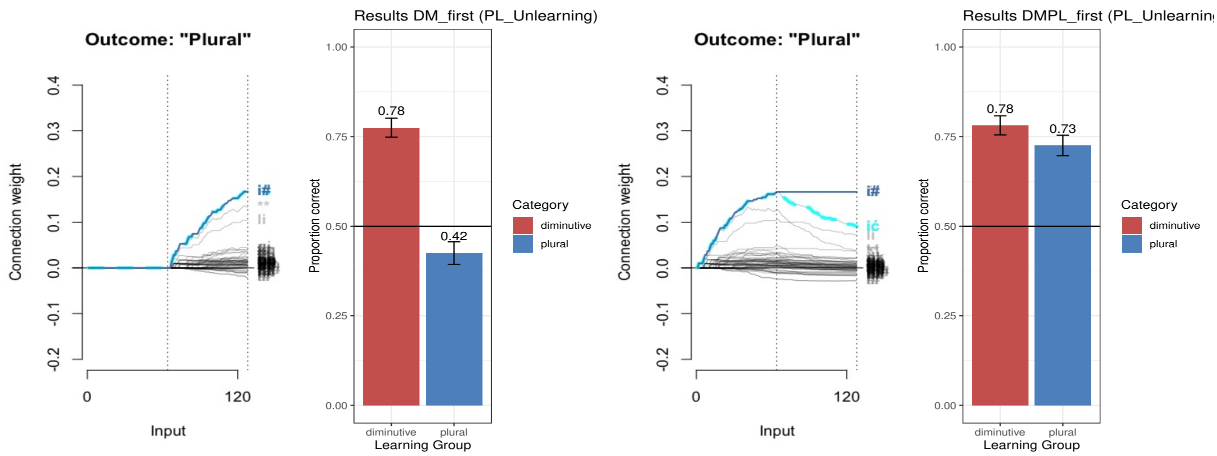


Figure 2: Results of the simulation (left most plots of the two pairs) and experimental results (right most plots) for PL_Unlearning

References

Hoppe, D. B., Hendriks, P., Ramscar, M., & van Rij, J. (2022). An exploration of error-driven learning in simple two-layer networks from a discriminative learning perspective. *Behavior Research Methods*, 1–31.

Nixon, J. S. (2020). Of mice and men: Speech sound acquisition as discriminative learning from prediction error, not just statistical tracking. *Cognition*, *197*, 104081.

Ramscar, M. (2021). A discriminative account of the learning, representation and processing of inflection systems. *Language, Cognition and Neuroscience*, 1–25.

Ramscar, M., Dye, M., & McCauley, S. M. (2013). Error and expectation in language learning: The curious absence of mouses in adult speech. *Language*, 760–793.

Ramscar, M., Yarlett, D., Dye, M., Denny, K., & Thorpe, K. (2010). The Effects of Feature-Label-Order and Their Implications for Symbolic Learning. *Cognitive Science*, *34*(6), 909–957.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement (A. H. Black & W. Prokasy, Eds.). *Classical conditioning II: current research and theory*, 64–99.

van Rij, J., & Hoppe, D. (2021). *edl: Toolbox for Error-Driven Learning Simulations with Two-Layer Networks* [R package version 1.0]. https://CRAN.R-project.org/package=edl

Evidence for a non-generic masculine generic in German

Dominic Schmitz[1], Viktoria Schneider[1], & Janina Esser[2]
[1]Heinrich-Heine-Universität Düsseldorf, [2]Universität zu Köln

Linguistic research has repeatedly demonstrated that masculine generics in German show a masculine bias (e.g. Gabriel et al., 2008; Gygax et al., 2008; Irmen & Kurovskaja, 2010; Koch, 2021; Misersky et al., 2019; Stahlberg & Sczesny, 2001). That is, grammatically masculine role-nouns such as *Anwalt* 'lawyer' can refer to men and women but may favour an interpretation in which mostly or only men are considered as potential referents (e.g. Misersky et al., 2019). However, research on this matter faces two major issues.

First, previous studies have used numerous of such masculine generics and their feminine counterparts to gain insights into their semantics without accounting for language external but potentially confounding influences such as stereotypicality. Second, the majority of studies finds the aforementioned masculine bias; however, very few studies offer a theoretical account of its underlying representations (e.g. Irmen & Linner, 2005). To this date, no investigation has been made to discover whether there is a connection between the male bias and the representation of masculine generics in the mental lexicon.

The present paper offers a solution to both issues. First, the role nouns under investigation are those for which stereotypicality ratings are available (Gabriel et al., 2008). Language external factors are incorporated in the analysis by such ratings. Second, using the general ideas of distributional semantics (Harris, 1954) as well as naïve discriminative learning (e.g. Baayen & Ramscar, 2015) and linear discriminative learning (e.g. Baayen et al., 2019) the underlying nature of masculine generics and counterparts is investigated. The proposed analysis aims at exploring the following question: How semantically similar are masculine generics and their explicitly masculine and feminine counterparts when taking into account stereotypicality?

The following method was employed to tackle this question. An 830,000 sentence (1.7 million words) corpus of contemporary German was created using the Leipzig Corpora Collection (Goldhahn et al., 2012). The corpus included 113 target word pairs which were based on the set of words provided by Gabriel et al. (2008). All target word occurrences were checked for their usage, i.e. whether they were generically or explicitly intended, and annotated accordingly. The corpus was then used to train semantic vectors based on the Rescorla-Wagner equation (Wagner & Rescorla, 1972) as implemented by naïve discriminative learning. Finally, the resulting semantic vectors were then used to train an implementation of linear discriminative learning.

Taking a closer look at the semantic vectors as created by the naïve discriminative learning algorithm, one finds that for the singular generic and explicit masculine role nouns are semantically closest with a mean cosine similarity value of approx. 0.98. Explicit masculine and feminine role nouns are less similar (approx. 0.94), and masculine generic and explicit feminine are least similar (0.93). These differences are highly significant (Mann-Whitney-Wilcoxon, p < 2.16e-16) and even more pronounced in the plural. To account for potential influences of stereotypicality, measures derived from the linear discriminative learning implementation were modelled by the stereotypicality measure as provided by Gabriel et al. (2008). As a result one finds that generic and explicit masculines are highly similar in terms of their semantic activation diversity and semantic neighbourhood size, while explicit feminines show higher activation diversities and denser neighbourhoods.

Our results indicate that there is a male bias in masculine generics in German as they exhibit highly similar semantic vectors as well as highly similar levels of semantic activation diversity and semantic neighbourhood size, even when controlled for their stereotypicality. Consequently, generic and explicit masculines show similar underlying representations, while the representations of explicit feminines are less similar. Thus, even though the use of masculine generics might be intended as generic, their resonance with the lexicon, that is more specifically their similarity with explicit masculines, leads to an overall biased association towards male referents.

References

Baayen, R. H., Chuang, Y.-Y., Shafaei-Bajestan, E., & Blevins, J. P. (2019). The discriminative Lexicon: A unified computational model for the lexicon and lexical processing in comprehension and production grounded not in (de)composition but in linear discriminative learning. *Complexity*, *2019*, 1–39. https://doi.org/10.1155/2019/4895891

Baayen, R. H., & Ramscar, M. (2015). Abstraction, storage and naive discriminative learning. *Handbook of Cognitive Linguistics*, *39*, 100–120. https://doi.org/10.1515/9783110292022-006

Gabriel, U., Gygax, P., Sarrasin, O., Garnham, A., & Oakhill, J. (2008). Au pairs are rarely male: Norms on the gender perception of role names across English, French, and German. *Behavior Research Methods*, *40*(1), 206–212. https://doi.org/10.3758/BRM.40.1.206

Goldhahn, D., Eckart, T., & Quasthoff, U. (2012). Building Large Monolingual Dictionaries at the Leipzig Corpora Collection: From 100 to 200 Languages. *Proceedings of the 8th International Language Resources and Evaluation (LREC'12)*.

Gygax, P., Gabriel, U., Sarrasin, O., Oakhill, J., & Garnham, A. (2008). Generically intended, but specifically interpreted: When beauticians, musicians, and mechanics are all men. *Language and Cognitive Processes*, *23*(3), 464–485. https://doi.org/10.1080/01690960701702035

Harris, Z. S. (1954). Distributional Structure. *WORD*, *10*(2–3), 146–162. https://doi.org/10.1080/00437956.1954.11659520

Irmen, L., & Kurovskaja, J. (2010). On the semantic content of grammatical gender and its impact on the representation of human referents. *Experimental Psychology*, *57*(5), 367–375. https://doi.org/10.1027/1618-3169/a000044

Irmen, L., & Linner, U. (2005). Die Repräsentation generisch maskuliner Personenbezeichnungen. *Zeitschrift Für Psychologie / Journal of Psychology*, *213*(3), 167–175. https://doi.org/10.1026/0044-3409.213.3.167

Koch, M. (2021). *Kognitive Effekte des generischen Maskulinums und genderneutraler Alternativen im Deutschen – eine empirische Untersuchung*. Technische Universität Braunschweig.

Misersky, J., Majid, A., & Snijders, T. M. (2019). Grammatical Gender in German Influences How Role-Nouns Are Interpreted: Evidence from ERPs. *Discourse Processes*, *56*(8), 643–654. https://doi.org/10.1080/0163853X.2018.1541382

Stahlberg, D., & Sczesny, S. (2001). Effekte des generischen Maskulinums und alternativer Sprachformen auf den gedanklichen Einbezug von Frauen. *Psychologische Rundschau*, *52*(3), 131–140. https://doi.org/10.1026//0033-3042.52.3.131

Wagner, A. R., & Rescorla, R. A. (1972). Inhibition in pavlovian conditioning: Application of a theory. In R. A. Boakes & M. S. Halliday (Eds.), *Inhibition and learning* (pp. 301–334). Academic Press Inc.