Title
Individuals differ cross-linguistically in cue weighting: A computational evaluation of cue-based retrieval in sentence processing

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors
Yadav, Himanshu
Smith, Garrett
Mertzen, Daniela
et al.

Publication Date
2022

Peer reviewed

Individuals differ cross-linguistically in cue weighting: A computational evaluation of cue-based retrieval in sentence processing

Himanshu Yadav (hyadav@uni-potsdam.de)
University of Potsdam, 14476 Potsdam, Germany

Garrett Smith (gasmith@uni-potsdam.de)
University of Potsdam, 14476 Potsdam, Germany

Daniela Mertzen (daniela.mertzen@uni-potsdam.de)
University of Potsdam, 14476 Potsdam, Germany

Relf Engbert (engbert@uni-potsdam.de)
University of Potsdam, 14476 Potsdam, Germany

Shravan Vasishth (vasishth@uni-potsdam.de)
University of Potsdam, 14476 Potsdam, Germany

Abstract

Cue-based retrieval theories of sentence processing assume that subject-verb dependencies are resolved through a content-addressable search in memory. The model assumes that multiple nouns with similar syntactic or semantic features increase dependency completion difficulty. English eyetracking data (reading) are consistent with model predictions; interestingly, a similar experiment with German–a language marking case overtly–suggests that only syntactic features affect dependency completion difficulty. Why would German show different behavior than English? Using a computational implementation of the cue-based retrieval model and model comparison using Bayes factors, we show that the reason is systematic variation at the individual-participant level: German participants overwhelmingly give higher weighting to syntactic cues over semantic cues, whereas English participants mostly give equal weighting to syntactic and semantic cues. The richer morphosyntax of German leads to syntactic cues being favoured; if such cues are largely absent (as in English) the parser relies on both cue types equally.

Keywords: Similarity-based interference; cue-based retrieval; individual differences

Introduction

Comprehending a sentence requires the reader to correctly figure out who did what to whom. This process of identifying the syntactic relations between words is called dependency completion. A well-established claim in sentence processing is that dependency completion between a verb and its associated subject is driven by a cue-based retrieval process (Lewis & Vasishth, 2005; McElree, 2000; Van Dyke, 2007). Under the cue-based retrieval account, the target noun is identified via a content-addressable search in memory based on feature specifications at the verb, such as [subject], called retrieval cues. When multiple nouns in memory match the retrieval cues, it is difficult to identify the target noun, which leads to a slowdown in retrieval times at the verb compared to a situation where only one noun matches the retrieval cues.

For example, in sentence (a) below, both the nouns *the resident* and *the neighbour* are in subject position, i.e., they both match the retrieval cue [subject], compared to sentence (b) where only one noun *the resident* matches the [subect] cue. The reading times at the verb *was complaining* are predicted to be slower in sentence (a) compared to sentence (b). This predicted effect is called *syntactic interference* (Van Dyke, 2007; Van Dyke & Lewis, 2003).

(a) ... the resident who said that the neighbour was dangerous was complaining ...

(b) ... the resident who was living near the dangerous neighbor was complaining ...

Similarly, when multiple nouns match the verb's semantic cues, such as [animate], they are assumed to cause *semantic interference* (Van Dyke, 2007). For example, in sentence (c) where both *the resident* and *the neighbour* are animate, the retrieval times at the verb are predicted to be slower than sentence (d), where only *the resident* is animate.

(c) ... the resident who said that the neighbour was dangerous was complaining ...

(d) ... the resident who said that the warehouse was dangerous was complaining ...

The predicted syntactic and semantic interference effects are consistently found in English reading studies (Van Dyke, 2007; Van Dyke & Lewis, 2003; Arnett & Wagers, 2017; Van Dyke & McElree, 2011). In a recent cross-linguistic study (Mertzen, Paape, Dillon, Engbert, & Vasishth, 2021), both syntactic and semantic interference were observed in English, but in German, only syntactic interference was observed at the verb (see Figure 1). Semantic interference was absent at the verb and appeared only later in the post-verbal region.

In sum, at the critical region (the verb phrase *was complaining*), syntactic interference predicted by the cue-based retrieval account is observed in both English and German, but semantic interference is observed only in English. The default assumption in cue-based retrieval models is that syntactic and semantic cues are used in the same way, so the magnitude of semantic and syntactic interference is predicted

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society*.

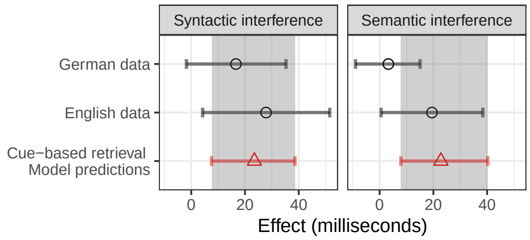

Figure 1: Syntactic and semantic interference predicted by a cue-based retrieval model (shaded areas; Lewis & Vasishth, 2005) compared with the observed effects in English and German from Mertzen et al. (2021); the effects are estimated from regression path durations at the verb. The error bars show 95% credible intervals of predicted or observed effects.

to be the same. The absence of semantic interference in German is, therefore, puzzling.

Mertzen et al. propose an explanation: a noun's syntactic cue is weighted higher than its semantic cue in German, while the two cues are weighted equally in English. The high weighting for syntactic cues in German could be due to the overt case marking on the nouns; it is possible that the overt case marking is highly reliable in identifying the grammatical functions of the nouns.

The cue weighting proposal is not entirely new. Several researchers have hypothesized that syntactic cues may be weighted more strongly over non-syntactic cues in processing antecedent-reflexive dependencies (Dillon, Mishler, Sloggett, & Phillips, 2013; Cunnings & Sturt, 2014; Kush, 2013; Parker & Phillips, 2017). A major limitation of these cue-weighting proposals is that they aim to explain the data averaged across all the participants. However, it is possible that individual differences in cue weighting exist. For example, a recent study on English (Yadav et al., 2021) showed that only one-third of the participants weigh syntactic cues more strongly over number cues in processing antecedent-reflexive dependencies. This result implies that the claim based on the average behavior holds only for a small subset of participants. This study demonstrates that the average behavior may mask theoretically important information which can only be revealed by modeling individual-level differences.

Modeling individual differences in syntactic and semantic interference in English and German might reveal a more nuanced picture of cue weighting differences among individuals and among the two language groups. For instance, it is possible that only a small subset of German participants, who have high weighting for the syntactic cue, is responsible for the absence of semantic interference in German. Therefore, the cue weighting hypothesis — that syntactic cues are weighted more strongly over semantic cues in German, but not in English — should be formulated for individual participants. The important questions to be asked are: (1) whether individuals differ in how they weight syntactic cues relative to semantic cues, and (2) whether individual German participants differ from individual English participants in cue weighting.

We test these questions by implementing two hierarchical models based on the Lewis and Vasishth (2005) cue-based retrieval model: (i) *the equal cue-weighting model*, which assumes that all the individuals have equal weights for syntactic and semantic cues, and (ii) *the varying cue-weighting model*, which assumes that individuals may differ in how highly they weight syntactic cues over semantic cues. The models are fitted to data from Mertzen et al. (2021) and then compared using Bayes factors (Rouder, Haaf, & Vandekerckhove, 2018).

The main finding is that there are cross-linguistic differences in individual-level cue-weighting: most German participants have higher weights for syntactic cues over semantic cues, while most English participants have equal weights for syntactic and semantic cues.

We first present the two individual difference models. Next, we quantify relative evidence for the two models and show the individual-level cue weighting estimates. We then discuss the broader implications of the work and conclude.

Two models of individual-level cue weighting

We implement two hierarchical models that differ in their assumption about the distribution of individual-level cue weighting. The models are implemented within the cue-based retrieval framework of Lewis and Vasishth (2005).

The Lewis and Vasishth (2005) model (see Engelmann, Jäger, & Vasishth, 2020, for the latest implementation) assumes that each noun phrase that matches a retrieval cue receives a certain amount of activation (see Figure 2). The total activation of a noun phrase i is given by

$$A_i = B_i + \sum_{j=1}^{n} W_j S_{ji} + \varepsilon_i \qquad (1)$$

where B_i is the baseline activation of the noun i determined by its past retrievals, and ε_i is Gaussian noise added to activation of the noun i, such that $\varepsilon_i \sim Normal(0, \sigma)$. The term $\sum_{j=1}^{n} W_j S_{ji}$ represents that the noun phrase i receives activation from all matching cues j depending on the associative strength S_{ji} between the cue j and the noun i, and the cue's weight W_j (Engelmann et al., 2020). The cue's weight is determined by a parameter called *cue weighting*. Cue weighting encodes the ratio of weights of syntactic cues and non-syntactic cues. Following Yadav et al. (2021), we assume that the cue weighting can have a value between 1 and 4, such that the cue weighting of 1 means equal weights for syntactic and semantic cues and the cue weighting of 4 means four times higher weight for the syntactic cue over semantic cues.

The Lewis and Vasishth model further assumes that a noun phrase with the highest activation gets retrieved for dependency completion. The retrieval time at the verb is determined by the activation level of the retrieved noun, A_i.

$$T = Fe^{-A_i} \qquad (2)$$

where the latency factor F reflects overall reading speed and may, inter alia, include lexical access time, motor response

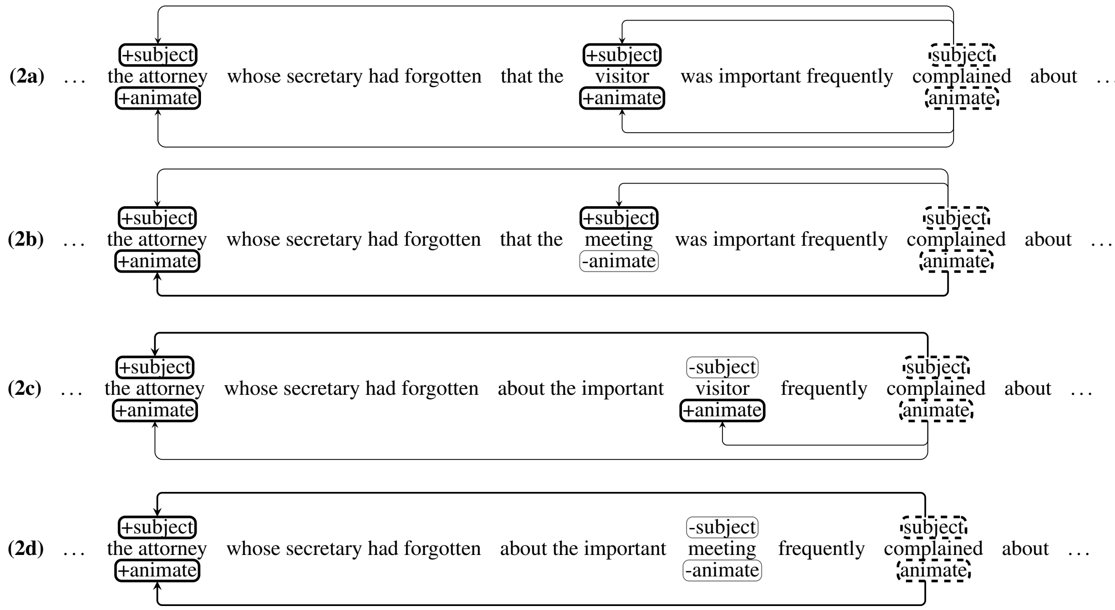

Figure 2: Activation received by the nouns based on cue-feature match, with a thick arrow denoting more activation spreading compared to a thin arrow. A dashed box represents a retrieval cue, a thick box represents a feature that matches a retrieval cue, and a thin box represents a feature that does not match a retrieval cue.

time, etc. The latency factor is commonly considered a free parameter in the model.

Figure 2 shows how activation spreads to each noun phrase in the four example conditions from Mertzen et al. (2021). In conditions (2a) and (2b), multiple nouns match the [subject] cue at the verb; as a result, the activation spread via the [subject] cue gets divided among these nouns. This is called the *fan effect* (Anderson et al., 2004; Schneider & Anderson, 2012). Due to the fan effect, the retrieval times at the verb in conditions (2a) and (2b) are predicted to be slower compared to conditions (2c) and (2d); this slowdown is referred to as syntactic interference. Similarly, due to the fan effect of the [+animate] feature in conditions (2a) and (2c), the retrieval times in (2a) and (2c) are predicted to be slower than in (2b) and (2d), which is called semantic interference.

Based on the equations 1 and 2, the model predicts syntactic and semantic interference, (X_{syn}, X_{sem}) as a function of cue weighting W and latency factor F,

$$(X_{syn}, X_{sem}) \sim Model(W, F) \tag{3}$$

The magnitude of both syntactic and semantic interference increases linearly with an increase in latency factor. But the cue weighting affects only semantic interference: the magnitude of semantic interference decreases with an increase in cue weighting. This is because with the increase in cue weighting, the semantic cue gets weaker, and consequently, the fan effect caused by the semantic cue gets weaker, which leads to the decrease in semantic interference.

We implement two hierarchical models that predict syntactic and semantic interference for each individual participant as a function of individual-level cue weighting and latency factor. The models make different assumptions about how the cue weighting varies among individuals, which we discuss next.

The equal cue-weighting model

The equal cue-weighting model assumes that all participants have equal weighting for syntactic and semantic cues.

Suppose that $(X_{syn_{j,g}}, X_{sem_{j,g}})$ represent syntactic and semantic interference effects for a participant j from language g.

$$(X_{syn_{j,g}}, X_{sem_{j,g}}) \sim Model(W_{j,g}, F_{j,g}) \tag{4}$$

where $W_{j,g}$ is the cue weighting and $F_{j,g}$ is the latency factor of the participant j of language g.

Under the equal cue-weighting model, all the participants regardless of their language have cue weighting equal to 1, i.e., they have equal weights for syntactic and semantic cues.

$$W_{j,g} = 1 \tag{5}$$

The individual-level latency factor $F_{j,g}$ is assumed to come from a normal distribution with population-level mean latency factor μ_{F_g} and population-level variance $\tau^2_{F_g}$ for language g:

$$F_{j,g} \sim Normal_{lb=0.05}(\mu_{F_g}, \tau^2_{F_g}) \tag{6}$$

where $lb = 0.05$ represent a lower bound of 0.05 on latency factor values. We choose this lower bound because a latency factor of less than 0.05 would generate unreasonably fast reading times for an individual (see Jäger, Engelmann, & Vasishth, 2017, for a meta-analysis of reading times).

The varying cue-weighting model

The varying cue-weighting model assumes that participants may differ in weighting of syntactic cues over semantic cues.

Suppose that $(X_{syn_{j,g}}, X_{sem_{j,g}})$ represents the syntactic and semantic interference effects for a participant j from the language g.

$$(X_{syn_{j,g}}, X_{sem_{j,g}}) \sim Model(W_{j,g}, F_{j,g}) \tag{7}$$

The individual-level latency factor $F_{j,g}$ is assumed to come from the same distribution as shown in Equation 6.

Under the varying cue-weighting model, the cue weighting for the participant j of language g, i.e., $W_{j,g}$ comes from a normal distribution with population-level mean cue weighting μ_{W_g} and between-participant variance $\tau^2_{W_g}$:

$$W_{j,g} \sim Normal_{lb=1,ub=4}(\mu_{W_g}, \tau^2_{W_g}) \tag{8}$$

where $lb = 1, ub = 4$ constrains the individual-level cue weighting to be between 1 and 4. A cue weighting of 1 means equal weights for syntactic and semantic cues and a cue weighting of 4 means 4 times higher weight for the syntactic cue.

The population-level cue weighting parameters, the mean cue weighting μ_{W_g} and between-participant variance $\tau^2_{W_g}$, are the main parameters that make the varying cue-weighting model different from the equal weighting model. A comparative evaluation of the two models can be sensitive to the priors on these population-level cue weighting parameters. Following the recommendation in Schad et al. (2021), we choose a range of priors on mean cue weighting and between-participant variance in cue weighting so that we can compare the models under different assumptions about the distribution of cue weighting in the populations.

For the population-level mean cue weighting μ_{W_g}, we specify the following prior:

$$\mu_{W_g} \sim Normal_{lb=1,ub=4}(1, \sigma_m) \tag{9}$$

where $\sigma_m \in \{0.05, 0.1, 0.5, 1\}$. The different values of σ_m express our assumptions about possible values of mean cue weighting. For example, $Normal_{lb=1,ub=4}(1, 0.05)$ represents that the mean cue weighting is retricted to be very close to 1, while $Normal_{lb=1,ub=4}(1, 1)$ represents that the mean cue weighting is allowed to be somewhere between 1 and 3.

For the between-participant variance in cue weighting $\tau^2_{W_g}$, we use an inverse-gamma prior.

$$\tau^2_{W_g} \sim InvGamma(1, scale) \tag{10}$$

where $scale \in \{0.005, 0.01, 0.05, 0.1, 0.5\}$. The different values of $scale$ express our assumptions about how much variation in cue weighting is allowed across individuals.

We fit these two models of individual-level cue weighting on data from Mertzen et al. (2021) and compute their marginal likelihoods given the data.

Model comparison

Mertzen et al. investigated both semantic and syntactic interference in a single design across two languages, English and German. From their dataset, we obtain shrunken estimates of individual-level syntactic and semantic interference for each participant as shown in Figure 3.[1]

We fit the equal cue weighting model and the varying cue-weighting model on the individual-level interference effects using hierarchical Approximate Bayesian Computation (Turner & Van Zandt, 2014; Sisson, Fan, & Beaumont, 2018) and obtained the marginal likelihoods for the each model given the data.

We then quantified the evidence for the varying cue-weighting model against the equal cue-weighting model using the Bayes factors (Rouder et al., 2018; Schönbrodt & Wagenmakers, 2018). The Bayes factor in favor of a model $\mathcal{M}_1$ compared to a model $\mathcal{M}_2$, i.e., BF_{12} is computed as the ratio of the marginal likelihoods of $\mathcal{M}_1$ and $\mathcal{M}_2$. The Bayes factor BF_{12} represents the extent to which the model $\mathcal{M}_1$ is more likely than $\mathcal{M}_2$ given the data. Following the convention from Jeffreys (1939/1998), a Bayes factor value of larger than 10 is interpreted as strong evidence in favor of $\mathcal{M}_1$ and a value between 3 and 10 is interpreted as moderate evidence in favor of $\mathcal{M}_1$.

Figure 4 shows the estimated Bayes factor under each prior assumption about the population-level cue weighting. We find that the Bayes factors are larger than 3 when the mean cue weighing is assumed to be very close to 1, suggesting moderate evidence in favor of the varying cue-weighting model. Under the assumption that the mean cue weighting could lie in the range of 1 to 2 or 1 to 3, the Bayes factors are larger than 10, indicating strong evidence in favor of the varying cue-weighting model.

Overall, the Bayes factors suggest moderate to strong evidence for the varying cue-weighting model compared to the equal cue-weighting model.

Individual-level cue weighting estimates

The model comparison shows evidence in favor of the assumption that individuals differ in cue weighting. But how do they differ? What is the distribution of individual-level cue weighting in English and German? We can answer this using individual-level cue weighting estimates from the varying cue-weighting model.

Figure 5 shows the estimated posterior distribution of cue-weighting for each individual participant from English and German. We find that 85% of the German participants have cue weighting larger than 2 meaning that 85% of the German participants give at least two times higher weights to syntactic cues over semantic cues. And, 84% of English participants have cue weighting of less than 1.5, which means that 84%

[1] To obtain individual-level interference effects, we fit a Bayesian hierarchical model with varying intercepts and slopes for participants and items, where regression path durations are the dependent variable and conditions (syntactic vs semantic, feature-match vs mismatch) are sum-coded predictors (Schad et al., 2020).

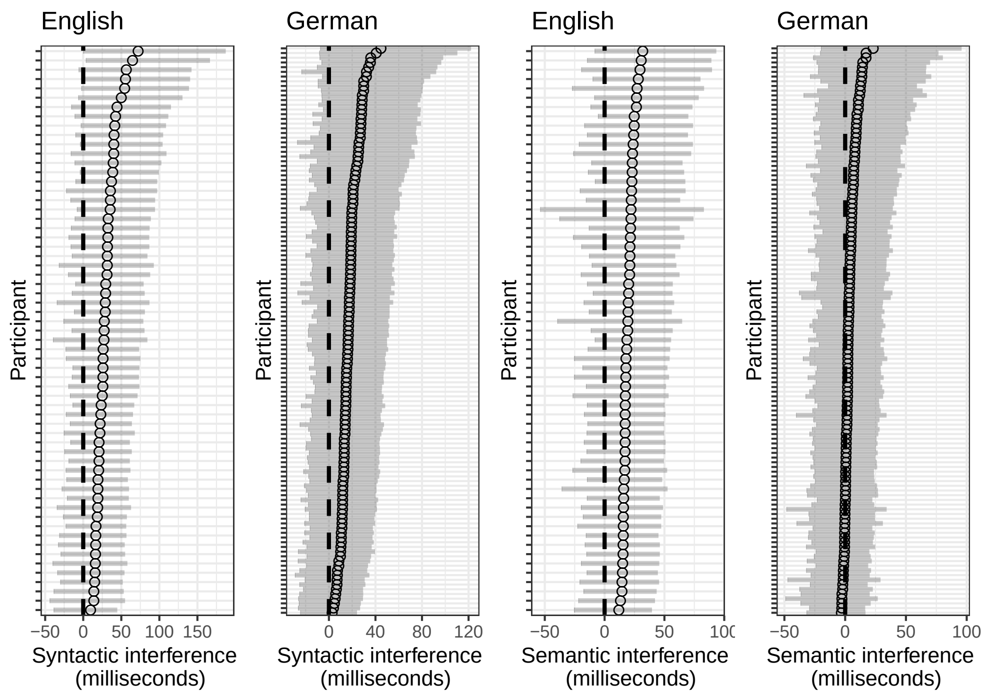

Figure 3: Individual-level syntactic and semantic interference effects from the Mertzen et al. (2021) data. Shown are the shrunken estimates from a Bayesian hierarchical model fit to regression path durations at the verb. English had 61 participants, German had 121 participants. The circles represent mean effects, the error bars represent 95% credible intervals of the effects.

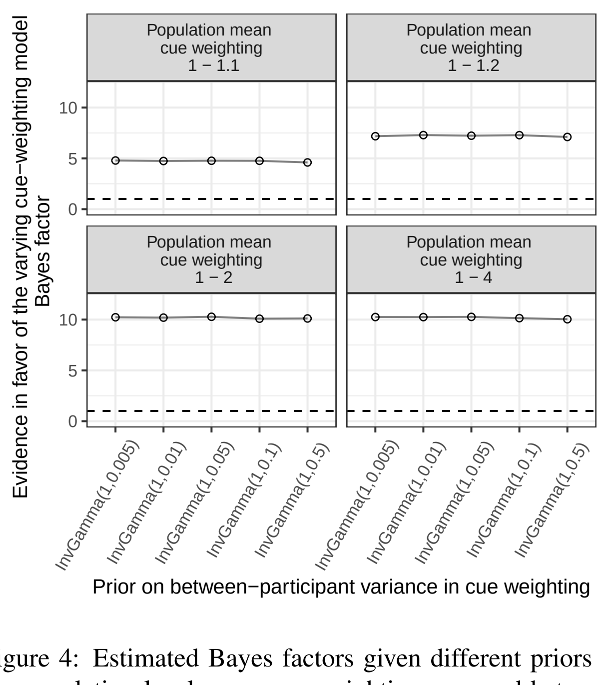

Figure 4: Estimated Bayes factors given different priors on the population-level mean cue weighting, μ_{W_g} and between-participant variance in cue weighting, $\tau^2_{W_g}$ (see Equation 8).

of English participants give approximately equal weights to syntactic and semantic cues.

In sum, the cue weighing estimates indicate that the most of the German participants have high weighting (> 2) for the syntactic cue, while the most of the English participants have equal weighting (≈ 1) for syntactic and semantic cues.

Discussion

Are syntactic and semantic retrieval cues weighted differently by English and German speakers? To answer this question, we implemented two hierarchical models, the equal cue-weighting model and the varying cue-weighting model. The equal cue-weighting model assumed that all English and German participants have equal weights for the syntactic and the semantic cues when retrieving a verb's subject from memory; the varying cue-weighting model assumed that individual participants can differ in how strongly they weight syntactic cues over semantic cues.

The models were evaluated on individual-level syntactic and semantic interference data from Mertzen et al. (2021). The model comparison and the model fits show that

1. There is moderate to strong evidence in favor of the varying cue-weighting model, suggesting that individuals vary in how they weight retrieval cues.

2. Most German participants give higher weights to syntactic cues over semantic cues, while most English participants give equal weights to syntactic and semantic cues.

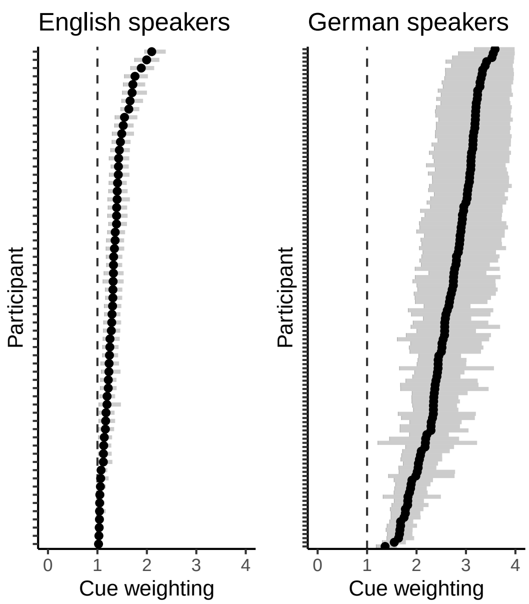

Figure 5: Individual-level cue weighting estimates from the varying cue-weigting model fitted to Mertzen et al. (2021) data. The solid circles represent mean cue weighting, the error bars represent 95% credible intervals of estimated values.

The results indicate that German speakers differ from English speakers in how they weight syntactic cues relative to non-syntactic cues. The conclusion is important for theories of sentence processing, because there is some independent support for the idea that the native speakers of a particular language may learn to use certain cues more strongly and reliably over the others (Dittmar et al., 2008; Sokolov, 1988; Bates et al., 1984). It is possible that German speakers weight the syntactic cues higher because the overt case marking in German is highly reliable in identifying the grammatical functions of the nouns in a sentence. For example, in the experimental items used in Mertzen et al. (2021), the grammatical role of every pre-verbal noun was identifiable either by (i) unambiguous case marking of the noun (if it was a masculine noun), or by (ii) the properties of its case-assigning head (verb or preposition).

A principled test of our conclusion would be in verifying whether the distribution of individual-level cue weighting in German and English is replicated in future experiments. If the inferred distribution — that most German speakers have high weighting for syntactic cues — holds for the language population, one would expect to see the same distribution of cue weighting in repeated experiments with larger samples of German participants. We plan to run a relatively large-sample-size study to test this prediction.

An interesting question that remains to be investigated is whether cue weighting is correlated with the general reading speed of an individual. There are reasons to believe that fast readers may weigh syntactic cues more strongly than slow readers (see Yadav et al., 2021). The strong weighting of syntactic cues in German speakers compared to English speakers might be associated with differences in their reading speed. Systematic experimental and modeling work is required to investigate the relationship between individual-level reading speed and cue weighting.

We have implemented only two models of individual differences in cue weighting, but one can explore other assumptions about how individuals vary in cue weighting. For example, it could be assumed that all German participants have a fixed cue weighting, which is different from English participants. Another assumption could be that only German participants vary in cue weighting while all English participants have equal cue weighting. It would be interesting to compare models under different assumptions about the distribution of individual-level cue weighting in German and English language populations.

A weakness of the current modeling work is that we do not have an independent measure of cue weighting for each individual; we can only infer it indirectly through reading times. The cue-weighting differences that are used to explain the observed individual differences in the data are estimated from the same data. It is possible that the individual-level cue weighting is overfitted to these data and that we may not get stable estimates of cue weighting for an individual in repeated experiments. A better approach would be to measure cue weighting independently for each participant on a separate processing task and then test the phenomenon of interest on the same group of participants. Using this approach, we can directly investigate whether the model can predict an individual's behavior based on their cue weighting. We plan to take this up in future work.

The current work reveals new insights about the constraints on processing subject-verb dependencies: The dependency between a verb and its associated subject is resolved via a cue-based retrieval process where the cues can be weighted differently by individuals depending on their native language. To our knowledge, this is the first investigation of cross-linguistic cue-weighting differences in a computational model of sentence comprehension. Our work contributes to understanding how different sources of linguistic information are employed during processing.

Acknowledgments

We thank the three anonymous reviewers for helpful suggestions. HY received funding from the Deutscher Akademischer Austauschdienst - DAAD, Programm ID: 57440921. SV received funding from Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Project-ID 317633480, SFB 1287.

References

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*(4), 1036–60.

Arnett, N., & Wagers, M. (2017). Subject encodings and retrieval interference. *Journal of Memory and Language*, *93*, 22–54.

Bates, E., MacWhinney, B., Caselli, C., Devescovi, A., Natale, F., & Venza, V. (1984). A cross-linguistic study of the development of sentence interpretation strategies. *Child development*, 341–354.

Cunnings, I., & Sturt, P. (2014). Coargumenthood and the processing of reflexives. *Journal of Memory and Language*, *75*, 117–139.

Dillon, B. W., Mishler, A., Sloggett, S., & Phillips, C. (2013). Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*, *69*, 85–103.

Dittmar, M., Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008). German children's comprehension of word order and case marking in causative sentences. *Child development*, *79*(4), 1152–1167.

Engelmann, F., Jäger, L. A., & Vasishth, S. (2020). The effect of prominence and cue association in retrieval processes: A computational account. *Cognitive Science*. Retrieved from 10.1111/cogs.12800

Jäger, L. A., Engelmann, F., & Vasishth, S. (2017). Similarity-based interference in sentence comprehension: Literature review and Bayesian meta-analysis. *Journal of Memory and Language*, *94*, 316-339.

Jeffreys, H. (1939/1998). *The theory of probability*. Oxford University Press.

Kush, D. (2013). *Respecting relations: Memory access and antecedent retrieval in incremental sentence processing*. Phd thesis, University of Maryland, College Park, MD.

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, *29*(3), 375–419.

McElree, B. (2000). Sentence comprehension is mediated by content-addressable memory structures. *Journal of Psycholinguistic Research*, *29*(2), 111–123.

Mertzen, D., Paape, D., Dillon, B., Engbert, R., & Vasishth, S. (2021). Syntactic and semantic interference in sentence comprehension: Support from English and German eye-tracking data.

Parker, D., & Phillips, C. (2017). Reflexive attraction in comprehension is selective. *Journal of Memory and Language*, *94*, 272–290.

Rouder, J. N., Haaf, J. M., & Vandekerckhove, J. (2018). Bayesian inference for psychology, part iv: Parameter estimation and bayes factors. *Psychonomic bulletin & review*, *25*(1), 102–113.

Schad, D. J., Nicenboim, B., Bürkner, P.-C., Betancourt, M., & Vasishth, S. (2021). *Workflow techniques for the robust use of bayes factors.* (draft)

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, *110*, 104038.

Schneider, D. W., & Anderson, J. R. (2012). Modeling fan effects on the time course of associative recognition. *Cognitive Psychology*, *64*(3), 127–160.

Schönbrodt, F. D., & Wagenmakers, E.-J. (2018). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic bulletin & review*, *25*(1), 128–142.

Sisson, S. A., Fan, Y., & Beaumont, M. (2018). *Handbook of approximate bayesian computation*. CRC Press.

Sokolov, J. L. (1988). Cue validity in hebrew sentence comprehension. *Journal of Child Language*, *15*(1), 129–155.

Turner, B. M., & Van Zandt, T. (2014). Hierarchical approximate bayesian computation. *Psychometrika*, *79*(2), 185–209.

Van Dyke, J. A. (2007). Interference effects from grammatically unavailable constituents during sentence processing. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *33*(2), 407–430.

Van Dyke, J. A., & Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language*, *49*, 285–316.

Van Dyke, J. A., & McElree, B. (2011). Cue-dependent interference in comprehension. *Journal of Memory and Language*, *65*(3), 247–263.

Yadav, H., Paape, D., Smith, G., Dillon, B. W., & Vasishth, S. (2021). Individual differences in cue weighting in sentence comprehension: An evaluation using Approximate Bayesian Computation. *Open Mind*. (Provisionally accepted)