Concreteness and Imageability Differentially Predict Judgments of Manual and Visual Similarity

Jonathan Wehnert, Dresden University of Technology, Germany Katharina von Kriegstein, Dresden University of Technology, Germany Brian Mathias, University of Aberdeen, United Kingdom

> The European Conference on Language Learning 2023 Official Conference Proceedings

Abstract

How we understand language is fundamentally shaped by our interactions with the world around us, according to modern theories of cognitive embodiment and predictive inference in perception. Sensorimotor experiences associated with specific words are reflected in those words' levels of concreteness and imageability, psycholinguistic constructs that are meant to capture the sensorimotor tangibility of word referents. Here, we investigate the relationship between word concreteness, imageability, and judgments of manual and visual similarity. Rather than examining features of individual words, we examined responses to word pairs consisting of German object nouns. Participants rated both the manual and visual similarity of each object word pair's referents. The ratings were correlated with intra-pair concreteness and imageability distances, based on norms previously acquired using a supervised learning algorithm (Köper & Schulte im Walde, 2016). Smaller concreteness distances between words in each pair correlated exclusively with higher manual similarity ratings, and smaller imageability distances correlated only with higher visual similarity ratings. This dissociation provides new evidence that concreteness and imageability can be thought of as distinct albeit somewhat related – features associated with different modalities. Whereas concreteness distance was more closely linked with manual object interactions, imageability distance was found to be more strongly associated with visual similarity. These findings may be of use to researchers designing stimulus sets for studies involving word pairs, such as linguistic priming or translation studies.

Keywords: Sensorimotor Semantics, Language Comprehension, Concreteness, Imageability



Introduction

Language provides a bridge between abstract thought and tangible reality. Hearing the word *apple*, for instance, may evoke its sweet taste, the crunchiness felt when biting into it, or even cultural contexts such as giving an apple to a schoolteacher. Such diverse associations highlight the complex interplay between language and our sensory experiences (Glenberg & Kaschak, 2002). Words are not merely symbolic placeholders; they carry a wealth of meaning derived from our interactions with the world (Clark, 1973; Lakoff & Johnson, 1980). The dynamic relationship between language and experience forms the core of what is often referred to as embodied cognition, which proposes that cognitive processes are fundamentally grounded in tangible experiences (Barsalou, 2020). This relationship is also fundamental to current theories of brain function such as the Bayesian brain hypothesis, which emphasizes the role of prior experience in generating predictions about incoming sensory information (Friston, 2012; Hipólito & Kirchhoff, 2023). A quintessential challenge for language learners, both native and non-native, is grasping the nuanced layers of meaning and associations that words carry, particularly concerning their concreteness and imageability (Baddeley et al., 1988; Paivio, 1971).

The concepts of concreteness and imageability have long held prominence in psycholinguistic research, given their implications for word recognition, memory, and language acquisition (Altarriba et al., 1999; Kroll & Merves, 1986). Concreteness in language refers to the extent to which words evoke tangible sensory and motor experiences (Barsalou, 1999; Glenberg & Kaschak, 2002). Words with high concreteness, such as bottle, evoke distinct sensorimotor experiences. Imageability, on the other hand, refers to the ease with which a word can evoke a clear mental image (Paivio, 1986; Schwanenflugel et al., 1988). Words with high imageability, such as *sunset*, typically elicit vivid visual imagery. When language learners encounter new words, their ability to understand and retain them is influenced by concreteness and imageability (Ellis & Beaton, 1993; Sadoski & Paivio, 2001). Highly concrete words, which are easily related to sensory experiences, are typically easier to remember and understand (de Groot & Keijzer, 2000). Similarly, words that evoke clear mental images can be more memorable because of the vividness of the imagery they invoke (Hayakawa et al., 2019). These findings parallel studies on language learning showing that multimodal forms of learning, including the viewing of relevant pictures and self-enactment of gestures, can benefit the retention of novel vocabulary (Mathias & von Kriegstein, 2023).

Although the concepts of concreteness and imageability are closely related, they differ in terms of their association with actions and the motor system. Concreteness often involves a direct or implicit link to actions, particularly in the context of object words. For example, the word *hammer* is concrete and can bring to mind not only the image of the tool but also the associated action of hammering. Manual interactions with objects can strongly influence their perceived concreteness (Engelen et al., 2011; Tillotson et al., 2008). In addition, the perception of action-related words can activate motor regions of the brain, evidencing the intertwining of language and motor systems (Pulvermüller et al., 2005). Imageability, however, appears to be more closely tied to the visual system. Even abstract words, such as *freedom*, can be highly imageable despite lacking a clear association with specific actions. Visual attributes of objects have been linked with their imageability (Cortese & Fugett, 2004), and high imageability words can trigger heightened activity in the visual cortex (D'Argembeau & Van der Linden, 2004).

This previous research leads to a yet untested prediction: Pairs of object words that are closely related in terms of concreteness might also be judged as more similar in terms of their action associations, because the things they refer to are likely to involve similar degrees of manual interaction. Moreover, pairs of object words that are more similar in terms of imageability may be perceived as having a higher degree of visual similarity. These predictions stem from the ideas that our comprehension of words is shaped by the motor and visual experiences linked to them (Barsalou, 2008), and that both manual interactions and visual attributes may influence how we understand and categorize objects (Borghi, 2005). Our study aimed to investigate the relationship between concreteness and manual similarity, as well as imageability and visual similarity. Participants rated object word pairs in terms of both manual and visual similarity. The ratings were then associated with variations in concreteness and imageability among the object pairs. We hypothesized that smaller distances in concreteness among object pairs would yield higher manual similarity ratings. Conversely, smaller distances in imageability among object pairs were expected to yield higher visual similarity ratings. Furthermore, we expected distances in concreteness to be more strongly associated with manual than visual similarity ratings and distances in imageability to be more strongly associated with visual than manual similarity ratings.

Method

Participants & Stimuli. The study included N = 17 native German speakers (9 female, 8 male, M age = 26 years, SD = 4.46, range: 20-34 years). Participants rated 115 German object word pairs in terms of visual similarity and manual similarity. The ratings were performed as part of a norming procedure for stimuli to be used in a future priming experiment, the results of which are not described here. Norms for the features of concreteness and imageability were derived from a corpus based on a supervised learning algorithm (Köper & Schulte im Walde, 2016). For 22 of the word pairs, one or both words were not part of the corpus of norms and those word pairs were therefore excluded from all analyses. This led to a total of 93 word pairs (e.g., Essstäbchen (chopsticks) – Bleistift (pencil)) consisting of 146 unique words being included in the study (while all word pairs were unique combinations of words, some words were used in multiple word pairs). All words were German object nouns. The 186 words (including duplicates) had a mean concreteness rating of 7.05 (SD = 0.77) and a mean imageability rating of 6.6 (SD = 0.8). Both ratings are given on a scale from 0 to 10. with higher ratings indicating higher concreteness and imageability, respectively (Köper & Schulte im Walde, 2016). Distance scores for concreteness and imageability were calculated as the absolute of the difference between the first and second words of each word pair.

Design & Procedure. In a visual similarity task, participants rated the German word pairs on a scale of 1 to 10 (from "not similar at all" to "very similar") in terms of visual similarity and, in a manual similarity task, they rated the same word pairs in terms of manual similarity. The order of the two tasks was counterbalanced across participants. The word pairs were shown in a randomized order in both tasks. The median, range, and interquartile range (IQR) of distance scores for concreteness and imageability, as well as of visual and manual similarity ratings are reported in **Table 1**. Both concreteness and imageability distance scores showed non-normal, right-skewed distributions, indicating relatively smaller distances for most word pairs. A Wilcoxon rank-sum test showed a significant difference between the mean manual (Mdn = 5, IQR = 6, range: 1 - 10) and visual (Mdn = 3, IQR = 4, range: 1 - 10) similarity ratings of word pairs (W = 5888, p < .001). This is not surprising, as the primary aim of the norming study was to find word pairs with greater manual similarity as compared to visual similarity for the future priming experiment. No significant difference was shown between

the concreteness (Mdn = 0.79, IQR = 0.88, range: 0.02 – 3.42) and imageability (Mdn = 0.63, IQR = 0.78, range: 0 – 3.1) distance scores (W = 4756.5, p = .24).

	Manual similarity	Visual similarity	Concreteness distance	Imageability distance
Median	5	3	0.79	0.63
Interquartile Range (IQR)	6	4	0.88	0.78
Range	1 – 10	1 – 10	0.02 - 3.42	0-3.1

Table 1. Median, interquartile range, and range for manual and visual similarity ratings, as well as concreteness and imageability distances.

Data Analysis. First, to test whether manual similarity was related to distances in concreteness and whether visual similarity was related to distances in imageability, we correlated the between-subject averages of the four dependent measures (manual similarity, visual similarity, concreteness distance, and imageability distance) with one another. Kendall's Tau-c was chosen as the correlation coefficient, as the data to be correlated differed in scale. Next, to see if these correlations differed significantly from one another, we calculated Kendall's Tau-c correlations individually for each participant and generated a linear mixed effects model of these correlations in R version 4.2.2 (R Core Team, 2022) using the 'lme4' (Bates et al., 2015) package. The full model included fixed effects of feature (concreteness distance, imageability distance) and task (manual ratings, visual ratings) as well as their interaction. The model's random effects structure was determined using backwards model selection (Barr et al., 2013), starting with a random intercept by participant and random slopes by participant for the independent factors of feature and task. We removed random effects terms one by one until the model no longer showed a singular fit. This resulted in the removal of the random slope by participant for the factor of feature. The final model therefore included a random intercept by participant and a random slope by participant for the independent factor of task. Significance testing was performed using Satterthwaite's method implemented with an alpha level of a = .05 (Kuznetsova et al., 2017). Post-hoc Tukey honestly significant difference (HSD) tests were used to conduct pairwise comparisons (Lenth et al., 2019).

Results

Our first hypothesis was that object word pairs that were more related in terms of concreteness would be rated as more manually similar. Confirming this hypothesis, a significant correlation between the concreteness distance and manual similarity measures was found ($\tau_c = -.19$, p = .007) across participants. Note that the negative character of this correlation indicates that manual similarity ratings were higher when distances in concreteness were smaller. We also predicted that object word pairs that were more related in terms of imageability would be rated as more visually similar. This prediction was also confirmed, as the imageability distance and visual similarity measures were also significantly negatively correlated ($\tau_c = -.19$, p = .006) across participants. Notably, no significant differences were found across participants between the manual similarity and imageability distance measures ($\tau_c = -.12$, p = .08), as well as between the visual similarity and concreteness distance measures ($\tau_c = -.09$, p = .21). These correlations are summarized in **Table 2**.

	Manual similarity	Visual similarity	
Concreteness distance	19**	09	
Imageability distance	12	19**	

Table 2. Group-level Kendall's Tau-c correlations between manual and visual similarityratings of object word pairs and concreteness and imageability distances for each word pair.There was a dissociation between concreteness and imageability distance in theirrelation to ratings of manual and visual similarity. **p < .01

Our last hypothesis was that distances in concreteness would be more strongly associated with manual than with visual similarity ratings, and that distances in imageability would be more strongly associated with visual than with manual similarity ratings. To assess this dissociation quantitatively, we calculated Kendall's Tau-c for each participant. Violin- and boxplots comparing the distribution of correlations across participants are shown in Figure 1. We then fitted a linear mixed effects model to these correlations with fixed effects for feature, task, and their interaction term, as well as a random intercept by participant and a random slope by participant for task. Critically, the model showed a significant interaction term between feature and task ($\beta = -0.15$, t = -7.68, p < .001, 95% CI [-0.19 -0.11]). Post-hoc pairwise comparisons were conducted using Tukey's HSD test. The post-hoc comparisons revealed that concreteness distances were correlated significantly more strongly with manual similarity (M = -0.15, 95% CI [-0.18 -0.13]) than with visual similarity (M = -0.07, 95% CI [-0.10 -0.05], p < .001). Conversely, imageability distances were correlated significantly more strongly with visual similarity (M = -0.17, 95% CI [-0.20 -0.15]) than with manual similarity (M = -0.11, 95% CI [-0.13, -0.08], p < .001). The mixed model also showed significant main effects of feature ($\beta = 0.05$, t = 3.58, p = .001, 95% CI [0.02 0.08]) and task ($\beta = 0.08$, t =5.49, p < .001, 95% CI [0.05 0.11]). These main effects indicate that, overall, correlations were more negative for concreteness distance scores than for imageability distance scores, and that correlations were more negative for manual than for visual similarity ratings. The full model results and pairwise comparison results are summarized in Table 3 and Table 4, respectively.

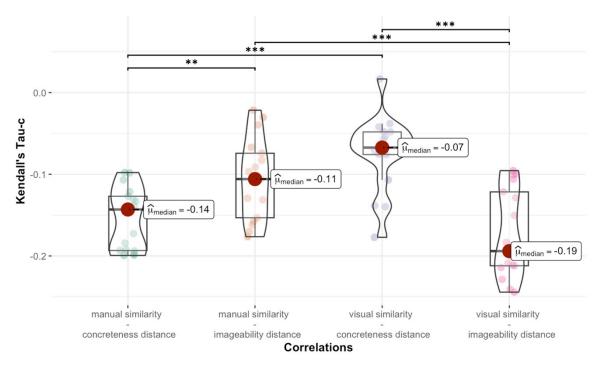


Figure 1. Combined violin- and boxplots of participant-level Kendall's Tau-c correlations between similarity ratings (manual, visual) and feature distances (concreteness, imageability). Concreteness distance correlated significantly more strongly with manual than visual similarity, and imageability distance was correlated significantly more strongly with visual than manual similarity. Note that a more negative Tau-c value indicates a stronger correlation between measures. **p < .01, ***p < .001

Fixed effects				Random effects					
	Estimate	SE	t	р	95% CI			Variance	SD
Intercept	15	.01	-15.10	<.001	18, 13	Participant	Intercept	.000	.02
Feature	.05	.01	3.58	.001**	.02, .08		Task	.001	.02
Task	.08	.01	5.49	<.001***	.05, .11				
Feature ×Task	15	.02	-7.68	<.001***	19, 11				

Table 3. Linear mixed effects model testing the effects of word pair feature (concreteness distance vs. imageability distance) and task (manual similarity vs. visual similarity) on the correlations between each of the four measures. **p < .01, ***p < .001.

Contrast	Estimate	SE	t ratio	р	95% CI	
concreteness-manual - imageability-manual	05	.01	-3.48	.007**	09	01
concreteness-manual - concreteness-visual	08	.02	-5.32	<.001***	12	04
concreteness-manual - imageability-visual	.02	.02	1.20	.63	02	.06
imageability-manual - concreteness-visual	03	.02	-2.11	.17	07	.01
imageability-manual - imageability-visual	.07	.02	4.42	<.001***	.03	.11
concreteness-visual - imageability-visual	.10	.01	7.06	<.001***	.06	.14

Table 4. Post-hoc Tukey's honestly significant difference (HSD) pairwise comparisonsdetailing the significant interaction between word features (concreteness, imageability)and task (manual, visual). **p < .01, ***p < .001

Discussion

This study aimed to characterize the relationship between sensorimotor aspects of word meanings and judgments of sensorimotor similarity. Specifically, we tested whether distances in concreteness and imageability features of object word pairs – rather than the features themselves – would be associated with ratings of manual and visual similarity. Indeed, concreteness distance of word pairs was found to be associated with manual similarity ratings, and imageability distance was associated with visual similarity ratings. Further, we observed a dissociation in the relationship between concreteness and imageability features and visual and motor modalities: Concreteness distance was found to be more associated with manual than visual similarity ratings and, conversely, imageability distance correlated more with visual than manual similarity ratings. These findings suggest that concreteness and imageability features can predict the perception of similarity between word referents. They also provide new evidence that concreteness and imageability reflect two differing psycholinguistic constructs, each related primarily to a different sensorimotor modality.

Concreteness and imageability have been found to predict the recognition and memorability of individual words (Ellis & Beaton, 1993; Li et al., 2020; Sadoski & Paivio, 2001; Soares et al., 2017). Our results extend these findings by demonstrating that concreteness and imageability also predict the extent to which two words are judged as similar or different with respect to their manual and visual associations. The dissociation between concreteness and imageability observed here is consistent with prior research showing that concreteness and imageability features engage distinct processing pathways. Whereas highly concrete words tend to elicit the simulation of sensory and motor experiences (Barsalou, 1999; Glenberg & Kaschak, 2002), words with high imageability evoke vivid visual imagery (Paivio, 1986; Schwanenflugel et al., 1988). This is in line with functional magnetic resonance imaging (fMRI) findings demonstrating that while both abstract and concrete nouns activate frontoparietal areas of the brain involved in action processes, concrete words do so more than abstract words (Del Maschio et al., 2022). Similarly, high imageability words can trigger heightened activity in the visual cortex (D'Argembeau & Van der Linden, 2004; Garbarini et al., 2020). These and the current findings are also consistent with theories emphasizing a key role of internal generative models in perception such as the multisensory predictive coding framework (MPCF; Mathias & von Kriegstein, 2023; Mayer et al., 2015) and the Bayesian brain hypothesis (Aitchison & Lengyel, 2017).

Another novel aspect of the current results is that the mere distance in concreteness and imageability between words (rather than the actual manual and visual properties of the words' referents, i.e., what actions each object affords and what each object looks like) predict ratings of manual and visual similarity, respectively. We speculate that, when both words of a word pair are of similar concreteness, they may activate the motor cortex to a similar degree. Analogously, two words of similar imageability may activate the visual cortex to similar degrees. It remains unknown however, how these similar levels of activation could translate into higher ratings of manual and visual similarity, given that one would expect the underlying patterns of activation to differ considerably between words referring to vastly different objects, even if those are similarly concrete and/or imageable. Future studies may shine a light on this question by employing multivariate analyses of the pattern of similarities between multiple word pairs and by looking at brain responses directly via neuroimaging methods.

This study had the primary aim of finding word pairs with greater manual than visual similarity to be used as stimuli in a subsequent priming experiment. For that reason, only object nouns were used as stimuli in the rating tasks, meaning that low concrete and imageable words were underrepresented in the current stimulus set. Future studies could improve on the stimulus selection by selecting word pairs more systematically to reflect broader spectra of concreteness and imageability, as well as manual and visual similarity. Likewise, they could arrange word pairs so that feature distances are distributed even more equally across high and low distance scores.

Conclusion

Using distance scores derived from automatically generated norms for features of concreteness and imageability (Köper & Schulte im Walde, 2016) and ratings for manual and visual similarity, the current study showed that word pair concreteness distance is associated with manual similarity ratings, and imageability distance is associated with visual similarity ratings. This finding is relevant to our understanding of the concepts of concreteness and imageability themselves. It may also be useful for researchers designing stimulus sets for studies involving word pairs, such as linguistic priming or translation studies.

References

- Aitchison, L., & Lengyel, M. (2017). With or without you: predictive coding and Bayesian inference in the brain. *Current Opinion in Neurobiology*, *46*, 219-227.
- Altarriba, J., Bauer, L. M., & Benvenuto, C. (1999). Concreteness, context availability, and imageability ratings and word associations for abstract, concrete, and emotion words. *Behavior Research Methods, Instruments, & Computers, 31*(4), 578-602.
- Baddeley, A. D., Papagno, C., & Vallar, G. (1988). When long-term learning depends on short-term storage. *Journal of Memory and Language*, 27(5), 586-595.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255-278.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22(4), 577-660.
- Barsalou, L. W. (2008). Grounded cognition. Annual Review of Psychology, 59, 617-645.
- Barsalou, L. W. (2020). Challenges and opportunities for grounding cognition. *Journal of Cognition*, *3*(1), 31.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), Article 1.
- Borghi, A. M. (2005). Object concepts and action. In D. Pecher & R. A. Zwaan (Eds.), *Grounding cognition: The role of perception and action in memory, language, and thinking* (pp. 8-34). Cambridge University Press.
- Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, *12*(4), 335-359.
- Cortese, M. J., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words. Behavior Research Methods, Instruments, & Computers, 36(3), 384-387.
- D'Argembeau, A., & Van der Linden, M. (2004). Influence of affective meaning on memory for contextual information. *Emotion*, 4(2), 173.
- De Groot, A. M., & Keijzer, R. (2000). What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning*, *50*(1), 1-56.
- Del Maschio, N., Fedeli, D., Garofalo, G., & Buccino, G. (2022). Evidence for the concreteness of abstract language: A meta-analysis of neuroimaging studies. *Brain Sciences*, *12*(1), 32.

- Ellis, R., & Beaton, A. (1993). Factors affecting the learning of foreign language vocabulary: Imagery keyword mediators and phonological short-term memory. *The Quarterly Journal of Experimental Psychology, 46*(3), 533-558.
- Engelen, J. A., Bouwmeester, S., de Bruin, A. B., & Zwaan, R. A. (2011). Perceptual simulation in developing language comprehension. *Journal of Experimental Child Psychology*, 110(4), 659-675.
- Friston, K. (2012). The history of the future of the Bayesian brain. *NeuroImage*, *62*(2), 1230-1233.
- Garbarini, F., Calzavarini, F., Diano, M., Biggio, M., Barbero, C., Radicioni, D. P., Geminiani, G., Sacco, K., & Marconi, D. (2020). Imageability effect on the functional brain activity during a naming to definition task. *Neuropsychologia*, 137, 107275.
- Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, 9(3), 558-565.
- Hayakawa, S., Bartolotti, J., van den Berg, A., & Marian, V. (2019). Language difficulty and prior learning influence foreign vocabulary acquisition. *Languages*, *5*(1), 2.
- Hipólito, I., & Kirchhoff, M. (2023). Breaking boundaries: The Bayesian Brain Hypothesis for perception and prediction. *Consciousness and Cognition*, 111, 103510.
- Köper, M., & Schulte im Walde, S. (2016). Automatically generated affective norms of abstractness, arousal, imageability and valence for 350 000 German lemmas. *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, 2595–2598.
- Kroll, J. F., & Merves, J. S. (1986). Lexical access for concrete and abstract words. *Journal* of Experimental Psychology: Learning, Memory, and Cognition, 12(1), 92.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). ImerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), Article 1.
- Lakoff, G., & Johnson, M. (1980). Metaphors we live by. University of Chicago Press.
- Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2019). Package "emmeans": Estimated marginal means, aka least-squares means. *Comprehensive R Archive Network*, 1–67.
- Li, D., Song, D., & Wang, T. (2020). Concreteness and imageability and their influences on Chinese two-character word recognition. *Reading and Writing*, *33*, 1443-1476.
- Mathias, B., & von Kriegstein, K. (2023). Enriched learning: behavior, brain, and computation. *Trends in Cognitive Sciences*, 27(1), 81-97.
- Mayer, K. M., Yildiz, I. B., Macedonia, M., & von Kriegstein, K. (2015). Visual and motor cortices differentially support the translation of foreign language words. *Current Biology*, 25(4), 530-535.

Paivio, A. (1971). Imagery and verbal processes. Holt, Rinehart, and Winston.

- Paivio, A. (1986). Mental representations: A dual coding approach. Oxford University Press.
- Pulvermüller, F., Hauk, O., Nikulin, V. V., & Ilmoniemi, R. J. (2005). Functional links between motor and language systems. *European Journal of Neuroscience*, 21(3), 793-797.
- R Core Team (2022). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. URL: https://www.R-project.org/.
- Sadoski, M., & Paivio, A. (2001). *Imagery and text: A dual coding theory of reading and writing*. Lawrence Erlbaum.
- Soares, A. P., Costa, A. S., Machado, J., Comesaña, M., & Oliveira, H. M. (2017). The Minho Word Pool: Norms for imageability, concreteness, and subjective frequency for 3,800 Portuguese words. *Behavior Research Methods*, 49(3), 1065–1081.
- Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language*, 27(5), 499-520.
- Tillotson, S. M., Siakaluk, P. D., & Pexman, P. M. (2008). Body-object interaction ratings for 1,618 monosyllabic nouns. Behavior Research Methods, 40(4), 1075-1078.

Contact email: jonathan.wehnert@tu-dresden.de