SEMANTIC TRANSPARENCY AFFECTS THE PHONETIC SIGNAL

Annika Schebesta[1], *Jessica Nieder*[2]
[1]*Universität Siegen*, [2]*Universität Passau*
annika.schebesta@uni-siegen.de

Abstract: Previous research has highlighted the impact of phonological factors and morphological structure on compound pronunciation, but the role of semantics in this interplay remains largely unexplored. Building on psycholinguistic findings that demonstrate semantic effects in compound processing, this study examines whether these effects extend to the production of English nominal triconstituent compounds. Using a state-of-the-art computational model to derive semantic transparency measures, we assessed the predictability of compound meanings from their constituents. Our regression model of experimental speech data revealed a notable duration difference between the first and second constituent and the third constituent for opaque compounds. For semantically transparent compounds, we find a shortening of the third constituent, diminishing the larger duration difference observed for less transparent compounds. These findings underscore the importance of semantic considerations in phonetic analyses of compound constituents, complementing prior research on morphological and phonological correlates of phonetic variation.

1 Introduction

Over the past years a growing body of linguistic evidence has shown that the phonetic signal of word forms interacts with their internal organization [e.g. 1, 2, 3, 4, 5, 6, 7], challenging earlier models of lexical processing that posit a strict separation of phonetics and morphology [8, 9]. More concretely, previous empirical research has yielded an effect of the presence of morphological boundaries on the duration and the quality of linguistic units.

Recent studies have demonstrated that phonetic factors can in fact influence different aspects of morphological structure in significant ways and vice versa. For example, research by Plag et al. [5] shows significant differences in acoustic duration between morphemic and non-morphemic word-final S in English. Additionally, word-final S varies according to the type of morphological boundary. For instance, plural S is significantly longer than 3rd person singular S and clitic S.Tomaschek et al. [7] reinforce these results by demonstrating that the discriminative capability of word-final S segments is reflected in their acoustic characteristics. They find that segments with higher discriminative capability tend to be articulated with longer durations, whereas segments with lower discriminative capability show overall shorter durations.

Similarly, Sproat and Fujimura [10] investigate the gradient velarization of /l/. Their results of articulatory measurements indicate that the categorical distinction of the allophones [l] and [ɫ] is inappropriate. Instead, the velarization of [l] appears to be gradient depending on the morphological boundary at which /l/ appears.

Compounds such as *blackbird* or *toothpaste* provide an intriguing testing ground for assessing the relationship of the phonetic signal and morphological structure [11]. Bell et al. [11] examined consonant durations at compound-internal boundaries in English compounds, finding that duration correlates positively with paradigmatic support and negatively with paradigmatic diversity.

The production and processing of larger morphological constructions has been the focus of Schebesta and Kunter [6] who investigate constituent durations in English left-branching, e.g. *healthcare law*, and right-branching, e.g. *corner drugstore*, triconstituent nominal compounds (NNN). Their statistical analysis reveals that the morphological organization of NNN compounds and the bigram frequency of two neighboring constituents enter a significant interaction. In particular, increasing N1N2 bigram frequencies have a significant shortening effect on N2 constituents in left-branching NNN, while N2 duration increase evidently in right-branching NNN. Similarly, high N2N3 bigram frequencies have a significantly stronger shortening effect on right-branching N3 constituents. The findings indicate that speakers make use of varying constituent durations in order to underpin the intended branching direction in cases where bigram frequencies interfere: The shortening of constituent durations is used to support the branching direction as determined by the morphological structure of the compounds.

Recent advances in computational linguistics and natural language processing have led to a number of new computational models that were proven useful for modelling various aspects of language [12, 13, 14, 15]. Crucially, in these models semantics plays a pivotal role, as it provides a deeper understanding of meaning that is essential for accurately capturing and processing language [16, 17, 18, 19, 20]. One notable example is the use of pre-trained models like BERT (Bidirectional Encoder Representations from Transformers) [15] which use contextual information from the sentential context in which words appear to generate rich semantic representations of words and phrases. These representations are not only crucial for improving performance on traditional NLP tasks but also allow for linguistically relevant applications such as the processing of compound words [16]. In the context of this study, the work by Buijtelaar and Pezzelle [16] demonstrates how measures such as semantic transparency, as derived from BERT models, can be effectively used to analyse and predict human semantic similarity judgements of compound words.

Inspired by recent work [20, 21], we apply the methodology of Buijtelaar and Pezzelle [16] to experimental data from a production study focused on NNN compounds by Schebesta [22]. Nieder et al. [20] used semantic measures taken from a computational model to successfully explain semantic effects in priming data that was collected in an earlier experimental study [23]. Similarly, Nieder et al. [21] detects semantic priming effects in auditory masked-priming data described in Ussishkin et al. [24] through computationally obtained semantic measures.

Following up on this, we aim to explore how semantic transparency influences the phonetic signal of English nominal triconstituent compounds. Our code and data is openly available at `https://osf.io/vfwes/`.

2 Experimental data

The data used in this study was elicited by Schebesta [22] in a production experiment with native speakers of North American English at the University of Alberta, Edmonton. The experimental study investigates the impact of branching direction, as determined by morphological structure, on the phonetic signal of English NNN compounds. It tests the prediction that morphologically embedded compound constituents show more phonetic reduction than free constituents.

In order to disentangle the effects of lexical bigram frequency and morphological structure on the duration of NNN compound constituents [6], the NNN in the production experiment were designed in such a way that the bigram frequencies of the neighboring nominal constituents were very low i.e., < 12 hits on COCA [25] so that an interaction of lexical bigram frequency and morphological structure is ruled out. In total, 25 W1W2 pairs e.g. *account service* were formed as the basis of NNN compounds. Each W1W2 pair was accompanied by a preceding or a following nominal constituent, yielding 50 different NNN e.g. *guest account service* and *account service assistant*. All NNN were spelled with spaces in order to prevent the orthographic form of the NNN from provoking a particular branching direction.

The 50 NNN compounds were embedded in short text passages that contained a context sentence and a carrier sentence. The morphological structure, that is, the branching direction of the NNN compounds was determined by the semantics of the context sentence, so that each NNN compound occurred once as left-branching and once as right-branching in the production experiment. The resulting 100 NNN compounds were produced by 42 participants (37 female, 5 male) who were recorded reading aloud the 100 text passages.

Using R [26], a linear mixed-effects regression model [27] predicted constituent durations and segment durations. The statistical model incorporated a number of phonological (number of phonemes, number of syllables, pitch measurements), lexical (phonological neighborhood density, lexical unigram frequency), and extra-linguistic (speech rate, number of repetitions) noise variables as well as two random intercepts for the speaker and the constituent.

The statistical analysis revealed that N1 and N2 constituent durations are equally long while N3 constituents are significantly longer, irrespective of branching direction. Following these results, the branching direction of the NNN compounds is not reflected in their phonetic signal, whereas the majority of the noise variables has the expected impact on constituent durations. This raises the question whether the semantics of the NNN compounds, that has shown an effect on processing in earlier studies, can explain the duration pattern of the compound constituents.

3 Method

3.1 Computational workflow

To investigate potential effects of the semantics of NNN on their acoustic signal, we implemented a computational workflow using IPython within the Jupyter Notebook environment. For acquiring word embeddings from BERT, we used the Hugging Face transformers library [28] to interact with the `BERTbase` model trained on English data [15]. Building upon the work of Buijtelaar and Pezzelle [16], we used their `nocontext_vector` function to obtain static meaning representations for each N of our NNN compounds. Notably, retrieving static embeddings from BERT in this way results in a meaning representation that is aggregated across multiple contexts.

In the next step, we calculated the semantic transparency (ST) of our compounds (c) following the methodology provided in Buijtelaar and Pezzelle [16]. ST quantifies how transparent the compound's meaning is based on the meanings of its constituents. Crucially, this calculation is based on cosine similarity between the left and the right constituent of the compound [see 16, for details on the original calculations]. Our study, however, involves triconstituent compounds, which differ in the number of constituents from the NN compounds analysed in their study. To accommodate for this difference, we extended the methodology to account for the additional constituent, ensuring that the semantic relationships among all three constituents of our compounds were appropriately captured and reflected in our ST calculations. To do so, we first created a combined vector for the embedded constituents by summing the vectors of the first embedded constituent (embed_1) and the second embedded constituent (embed_2) and taking the average of this combined vector. For ST, the cosine similarity between this embedded compound vector and the free constituent was calculated. The formula for semantic transparency as adapted from Buijtelaar and Pezzelle [16] is given below.

$$\mathrm{ST}(c) = \frac{\mathbf{v}_{\text{free}} \cdot \left(\frac{\mathbf{v}_{embed_1} + \mathbf{v}_{embed_2}}{2}\right)}{\left\|\mathbf{v}_{\text{free}}\right\| \cdot \left\|\frac{\mathbf{v}_{embed_1} + \mathbf{v}_{embed_2}}{2}\right\|}$$

The ST measure was then added to the experimental data and subsequently given as input to the R environment [26] for further statistical modelling.

3.2 Statistical workflow

For our statistical modelling procedure, we opted for a linear mixed-effect regression model using the R package `lme4` [27]. The preliminary regression model is provided in (1). It contains the dependent variable DURATION which corresponds to the constituent durations as annotated in Praat [29]. The statistical analysis includes an interaction of MEMBER × SEMANTIC TRANSPARENCY × BRANCHING, where MEMBER has the three levels `N1`, `N2` and `N3` corresponding to the three compound constituents, and BRANCHING contains the levels `left-branching` and `right-branching`. Equipped in this way, the regression model is able to predict any effect of SEMANTIC TRANSPARENCY on individual members of the NNN compounds or on members from NNN with a particular branching direction.

The regression model is complemented with a set of noise variables, such as LGUNIGRAMFREQ (the frequency of individual constituents log-transformed to the base 10), the SPEECH RATE at which a NNN compound was produced, and REPETITION which indicates the number of repetitions of the contained W1W2 pairs. Moreover, three Principal Component Analyses [30] were performed with (i) six different pitch measurements (generated with Praat), (ii) four predictors that inform about the phonological form and neighborhood of constituents, and (iii) two predictors that inform about the phonological form of NNN compounds ((ii) and (iii) from the *English Lexicon Project* [31]) in order to minimize the risk of collinearity. The most informative of the resulting principal components are PC1PITCH, PC2PITCH and PC3PITCH (i), PC1PHON and PC2PHON (ii), and PC1NNNPHON (iii), all of which are included in the regression model. In addition, the random intercepts (1 | SPEAKER) and (1 | CONSTITUENT) are included in the statistical analysis to account for speaker-specific variation and a potential influence of features of individual constituents.

(1) DURATION ~ (1 | SPEAKER) + (1 | CONSTITUENT)
+ MEMBER × SEMANTIC TRANSPARENCY × BRANCHING
+ LGUNIGRAMFREQ
+ SPEECH RATE
+ REPETITION
+ PC1PITCH + PC2PITCH + PC3PITCH
+ PC1PHON + PC2PHON
+ PC1NNNPHON

As the distribution of residuals of the preliminary model did not meet the normality assumption of linear regression, DURATION was Box-Cox-transformed [32] using an exponent of $\lambda = 0.02$. Additionally, 29 outliers from a total of 10,710 observations from 3,573 NNN compounds were excluded from the data set ($\triangleq$ 0.27%). With the modified dependent variable and the reduced data set, the final regression model was tested for correlated predictors using the `vif()` function from the `car` package [33], which did not reveal any collinearity. The regression model was not further reduced in order to minimize the risk of overfitting and Type I errors [34, 35].

4 Results

The results from our linear mixed-effect regression model ($R^2_{marg.}$ 0.562, $R^2_{cond.}$ 0.797, AIC -96542.034) show that N1 and N2 constituents are significantly shorter than N3 constituents, irrespective of branching direction. The noise variables affect constituent durations as expected: phonologically longer constituents with fewer phonological neighbors and constituents with a pitch accent display longer durations, and repetitions of constituents and a higher speech rate lead to shorter constituent durations.

The interaction of MEMBER × SEMANTIC TRANSPARENCY reveals that only N3 constituents are significantly affected by SEMANTIC TRANSPARENCY. As can be observed in the left and middle panel of Figure 1, N1 and N2 durations (measured in seconds) remain the same regardless of increasing SEMANTIC TRANSPARENCY. For N3 durations, displayed in the right panel, we observe a significant facilitatory effect for greater values of semantic transparency, i.e. more transparent NNN.

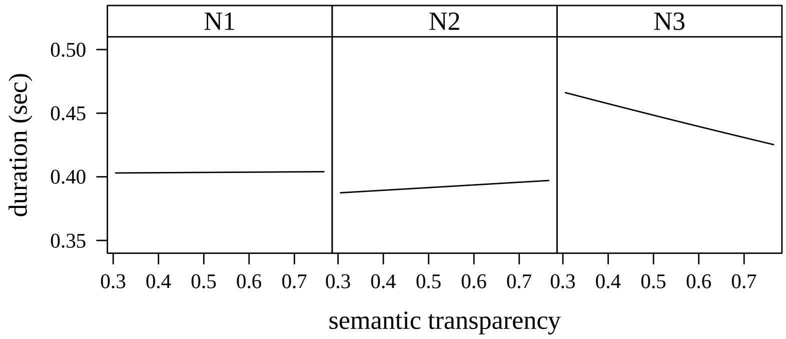

Figure 1 – Constituent duration by semantic transparency and member

5 Discussion

In this study, we tested the effect of semantics on the phonetic signal of NNN compounds. Previous work on NNN compounds has mainly focused on an interaction of morphological structure and the acoustic signal, leaving possible effects of semantics aside.

Recent work has shown the potential of computational models in detecting semantic effects in post-hoc analyses of experimental data [20, 21]. Following this line of research, we extracted distributional meaning representations, i.e. word embedding vectors, for our NNN compounds from a large language model. Using these embeddings, we calculated the semantic transparency of the compounds by assessing the meaning difference between the embedded compound vector and the vector of the free constituent through a cosine similarity calculation [16].

In our statistical model, we observed an effect for opaque vs. transparent compounds: In more opaque NNN, the first and second constituents, N1 and N2, were significantly shorter than the third constituent N3. While this duration pattern remained for semantically transparent NNN, the overall difference in duration between N1N2 and N3 became less pronounced as N3 constituent durations decreased. We interpret this effect as follows. In semantically opaque compounds, speakers indicate the opaqueness through a lengthening of the third constituent, emphasizing the difference in meaning between the embedded compound and the free constituent. In semantically transparent compounds, on the other hand, speakers do not highlight the relationship of the constituents as their meaning emerges naturally from the whole NNN. Thus, previously observed duration differences for opaque compounds become less pronounced as N3 durations decrease.

Our results provide evidence that the semantic relationship of compound constituents needs to be taken into account for studies focusing on an interaction of morphological structure and the acoustic signal. We show that including measures such as semantic transparency in a post-hoc analysis of experimental data can reveal previously unseen patterns. This can help in further assess and enrich data seeking to explain the morphology-phonetics interaction to gain a more comprehensive understanding of language processing.

References

[1] BEN HEDIA, S. and I. PLAG: *Gemination and degemination in English prefixation: Phonetic evidence for morphological organization*. *Journal of Phonetics*, 62, pp. 34–49, 2017. doi:https://doi.org/10.1016/j.wocn.2017.02.002. URL `https://www.sciencedirect.com/science/article/pii/S0095447017300372`.

[2] BLAZEJ, L. J. and A. M. COHEN-GOLDBERG: *Can we hear morphological complexity before words are complex?* *Journal of Experimental Psychology: Human Perception and Performance*, 41(1), pp. 50–68, 2015. doi:10.1037/a0038509. URL `https://doi.org/10.1037/a0038509`.

[3] HAY, J. and I. PLAG: *What Constrains Possible Suffix Combinations? On the Interaction of Grammatical and Processing Restrictions in Derivational Morphology*. *Natural Language & Linguistic Theory*, 22(3), pp. 565–596, 2004. doi:10.1023/B:NALA.0000027679.63308.89. URL `https://doi.org/10.1023/B:NALA.0000027679.63308.89`.

[4] KUNTER, G. and I. PLAG: *Morphological embedding and phonetic reduction: the case of triconstituent compounds*. *Morphology*, 26, pp. 201–227, 2016. doi:10.1007/s11525-016-9284-5.

[5] PLAG, I., J. HOMANN, and G. KUNTER: *Homophony and morphology: The acoustics of word-final S in English*. *Journal of Linguistics*, 53(1), p. 181–216, 2017. doi:10.1017/S0022226715000183.

[6] SCHEBESTA, A. and G. KUNTER: *Constituent durations in English NNN compounds: A case of strategic speaker behavior?* *Journal of Phonetics*, 94, 2022. doi:10.1016/j.wocn.2022.101164.

[7] TOMASCHEK, F., I. PLAG, M. ERNESTUS, and R. H. BAAYEN: *Phonetic effects of morphology and context: Modeling the duration of word-final S in English with naïve discriminative learning*. *Journal of Linguistics*, 57(1), pp. 123–161, 2021. doi:10.1017/S0022226719000203.

[8] BERMÚDEZ-OTERO, R.: *The architecture of grammar and the division of labor in exponence*. In *The Morphology and Phonology of Exponence*. Oxford University Press, 2012. doi:10.1093/acprof:oso/9780199573721.003.0002. URL `https://doi.org/10.1093/acprof:oso/9780199573721.003.0002`. `https://academic.oup.com/book/0/chapter/270627764/chapter-pdf/39032501/acprof-9780199573721-chapter-2.pdf`.

[9] LEVELT, W. J. M., A. ROELOFS, and A. S. MEYER: *A theory of lexical access in speech production*. *Behavioral and Brain Sciences*, 22(1), pp. 1–38, 1999. doi:10.1017/S0140525X99001776.

[10] SPROAT, R. and O. FUJIMURA: *Allophonic Variation in English /l/ and its Implications for Phonetic Implementation*. *Journal of Phonetics*, 21, pp. 291–311, 1993.

[11] BELL, M. J., S. BEN HEDIA, and I. PLAG: *How morphological structure affects phonetic realisation in English compound nouns*. *Morphology*, 31, pp. 87–120, 2021. doi:10.1007/s11525-020-09346-6.

[12] BAAYEN, R. H., Y.-Y. CHUANG, E. SHAFAEI-BAJESTAN, and J. P. BLEVINS: *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.

[13] JOULIN, A., E. GRAVE, P. BOJANOWSKI, M. DOUZE, H. JÉGOU, and T. MIKOLOV: *Fasttext.zip: Compressing text classification models*. *arXiv*, 2016. URL `https://arxiv.org/abs/1612.03651`.

[14] JOULIN, A., E. GRAVE, P. BOJANOWSKI, and T. MIKOLOV: *Bag of tricks for efficient text classification*. *arXiv*, 2016. URL `https://arxiv.org/abs/1607.01759`. `1607.01759`.

[15] DEVLIN, J., M.-W. CHANG, K. LEE, and K. TOUTANOVA: *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. In J. BURSTEIN, C. DORAN, and T. SOLORIO (eds.), *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pp. 4171–4186. Association for Computational Linguistics, Minneapolis, Minnesota, 2019. doi:10.18653/v1/N19-1423.

[16] BUIJTELAAR, L. and S. PEZZELLE: *A Psycholinguistic Analysis of BERT's Representations of Compounds*. In A. VLACHOS and I. AUGENSTEIN (eds.), *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, pp. 2230–2241. Association for Computational Linguistics, Dubrovnik, Croatia, 2023. doi:10.18653/v1/2023.eacl-main.163.

[17] FEIFEI, X., Z. SHUTING, and T. YU: *BERT-based Siamese Network for Semantic Similarity*. *Journal of Physics: Conference Series*, 1684(1), p. 012074, 2020. doi:10.1088/1742-6596/1684/1/012074. URL `https://dx.doi.org/10.1088/1742-6596/1684/1/012074`.

[18] HEITMEIER, M., Y.-Y. CHUANG, and R. H. BAAYEN: *Modeling morphology with linear discriminative learning: considerations and design choices*. *Frontiers in Psychology*, 12, 2021. doi:10.3389/fpsyg.2021.720713.

[19] MIKOLOV, T., E. GRAVE, P. BOJANOWSKI, C. PUHRSCH, and A. JOULIN: *Advances in pre-training distributed word representations*. In *Proceedings of the International Conference on Language Resources and Evaluation (LREC 2018)*. 2018.

[20] NIEDER, J., Y.-Y. CHUANG, R. VAN DE VIJVER, and R. H. BAAYEN: *A Discriminative Lexicon Approach to Word Comprehension, Production and Processing: Maltese Plurals*. *Language*, 99(2), pp. 242–274, 2023.

[21] NIEDER, J., R. VAN DE VIJVER, and A. USSISHKIN: *Emerging roots: Investigating early access to meaning in Maltese auditory word recognition*. *PsyArXiv*, 2024. doi:https://doi.org/10.31234/osf.io/hwumf.

[22] SCHEBESTA, A.: *NNN compounds in English: Investigating the interface of morphology, lexical frequency and the phonetic signal*. Ph.D. thesis, Universität Siegen, 2024. [Unpublished doctoral dissertation].

[23] NIEDER, J., R. VAN DE VIJVER, and H. MITTERER: *Priming Maltese plurals: Representation of sound and broken plurals in the mental lexicon*. *The Mental Lexicon*, 16(1), pp. 69–97, 2021.

[24] USSISHKIN, A., C. R. DAWSON, A. WEDEL, and K. SCHLUTER: *Auditory masked priming in Maltese spoken word recognition*. *Language, Cognition and Neuroscience*, 30(9), pp. 1096–1115, 2015.

[25] DAVIES, M.: *The Corpus of Contemporary American English (COCA)*. `https://www.english-corpora.org/coca/`, 2008.

[26] R CORE TEAM: *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2023. URL `https://www.R-project.org/`.

[27] BATES, D., M. MÄCHLER, B. BOLKER, and S. WALKER: *Fitting linear mixed-effects models using lme4*. *Journal of Statistical Software*, 67(1), pp. 1–48, 2015. doi:10.18637/jss.v067.i01.

[28] WOLF, T., L. DEBUT, V. SANH, J. CHAUMOND, C. DELANGUE, A. MOI, P. CISTAC, T. RAULT, R. LOUF, M. FUNTOWICZ, J. DAVISON, S. SHLEIFER, P. VON PLATEN, C. MA, Y. JERNITE, J. PLU, C. XU, T. L. SCAO, S. GUGGER, M. DRAME, Q. LHOEST, and A. M. RUSH: *Huggingface's transformers: State-of-the-art natural language processing*. 2020. `1910.03771`.

[29] BOERSMA, P. and D. WEENINK: *Praat: Doing phonetics by computer*. `http://www.praat.org/`, 2017.

[30] JAMES, G., D. WITTEN, T. HASTIE, and R. TIBSHIRANI: *An introduction to statistical learning with applications in R*. Springer, New York, 6th printing 2015 edn., 2013.

[31] BALOTA, D. A., M. J. YAP, M. J. CORTESE, K. A. HUTCHISON, B. KESSLER, B. LOFTIS, J. H. NEELY, D. L. NELSON, G. B. SIMPSON, and R. TREIMAN: *The English Lexicon Project*. *Behavior Research Methods*, 39(3), pp. 445–459, 2007.

[32] BOX, G. and D. R. COX: *An analysis of transformations*. *Journal of the Royal Statistical Society*, 26(2), pp. 211–256, 1964.

[33] FOX, J. and S. WEISBERG: *An R Companion to Applied Regression*. Sage, Thousand Oaks CA, third edn., 2019. URL `https://socialsciences.mcmaster.ca/jfox/Books/Companion/`.

[34] HARRELL, F. E.: *Regression modeling strategies: With applications to linear models, logistic regression, and survival analysis*. Springer eBook Collection Mathematics and Statistics. Springer, New York, NY, 2001. doi:10.1007/978-1-4757-3462-1.

[35] WINTER, B.: *Statistics for linguists: An introduction using R*. Routledge Taylor & Francis Group, New York and London, 2020. doi:10.4324/9781315165547. URL `https://www.taylorfrancis.com/books/9781315165547`.