**Semantics and pitch contour predict one another in English compounds**

Dominic Schmitz[1], Melanie J. Bell[2] and Ingo Plag[1]
*[1]Heinrich Heine University Düsseldorf, [2]Anglia Ruskin University*
{dominic.schmitz, ingo.plag}@uni-duesseldorf.de, melanie.bell@aru.ac.uk

Although prominence in English noun-noun compounds has been described in terms of a binary left-right stress distinction, the actual prosodic realisation of these expressions is far richer and more variable than such categorical labels suggest. Existing work has shown that many compounds allow multiple stress patterns across speakers and contexts [1, 2], but the precise phonetic basis of this variability and its theoretical interpretation remain unresolved. There is a substantial body of work linking compound stress to various semantic and lexical factors, including, e.g., semantic category, specificity and lexicalisation of the compound, semantic category and informativity of the constituents, and semantic relations between the constituents [1-5], but the nature of these connections has remained difficult to capture. The present study approaches this issue by analysing compound stress directly in the acoustic signal and by modelling how listeners could learn stress patterns from experience.

Using 971 compound tokens from the Boston University Radio Speech Corpus [6], we derived continuous pitch trajectories from raw pitch data using generalised additive models that smooth micro-prosodic variation, interpolate voiceless regions, and account for speaker-specific pitch realisation. Figure 1 illustrates the difference between raw pitch data and GAM-based pitch contours. K-means clustering of the GAM-based contours consistently yields two robust contour clusters corresponding to left- and right-stressed patterns, confirming the phonetic reality of the distinction while revealing considerable within-cluster variation (Figure 2). The cluster structure mirrors established phonetic findings [2, 7] but also demonstrates that what is typically labelled “right stress” encompasses a range of shapes rather than a single canonical contour (Figure 3).

To investigate how such stress patterns could be acquired and used in comprehension and production, we implemented a linear discriminative learning model [8] to map between pitch contours and context-dependent semantic embeddings computed using the large language model BERT [9]. In this framework, pitch patterns emerge not as abstract categories but as the outcome of statistical associations learned from experience between acoustic trajectories and the meanings related to them. The model successfully predicted compounds’ semantic embeddings from pitch contours, and vice versa, performing well above chance in both directions. The model requires no explicit stress labels, no categorical stress features, and no manually coded semantic relations or categories: stress is encoded and decoded through its distributional covariation with meaning in context.

Taken together, our results suggest the following view of stress in English compounds. First, stress as encoded in the continuous shape of the pitch contour falls into two broad clusters without collapsing into rigid categories. Second, the variability observed across tokens and speakers is not noise but an interpretable signal that reflects both production differences and the contextualised meanings speakers express. Third, stress patterns can be learned in comprehension and production through discriminative mappings between acoustic trajectories and semantics, offering a dynamic alternative to categorical stress rules. In sum, the study provides empirical evidence for a view of compound stress grounded in the structure of the speech signal and its association with meaning in context, rather than in abstract stress categories or discrete semantic primitives.

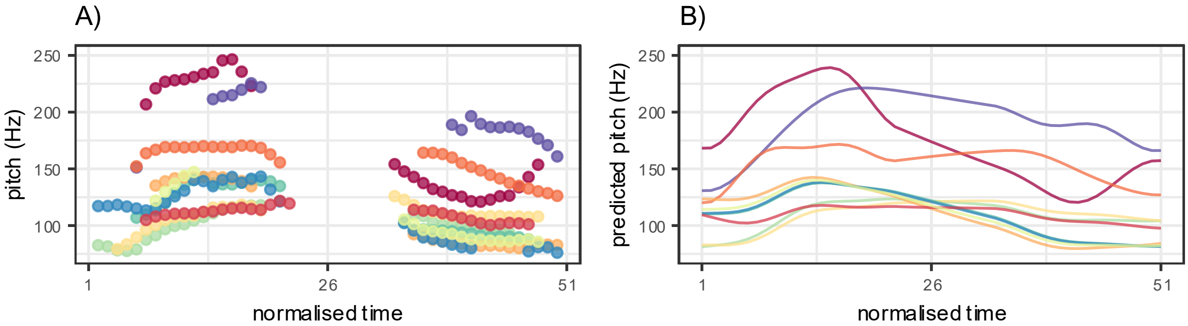


Figure 1: Panel A) shows the raw pitch points of single tokens of the compound *health care* produced by different speakers in different contexts; panel B) shows the individual pitch contours as predicted by the GAM; time is normalised

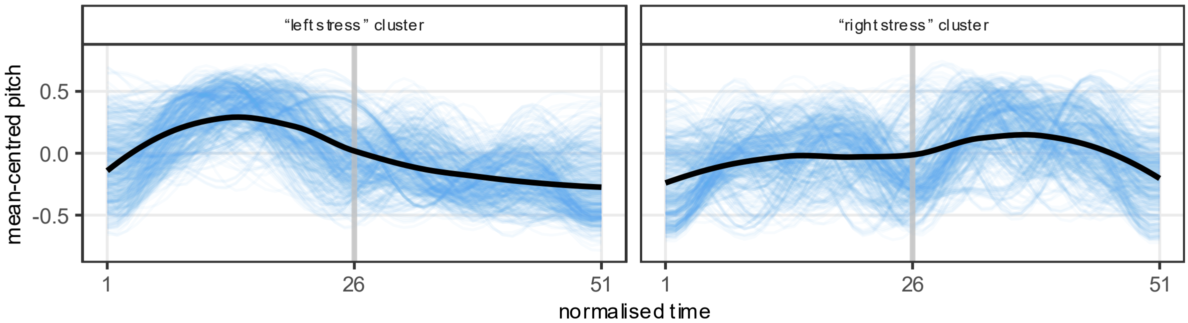


Figure 2: The identified pitch contour clusters, with a non-linear average smooth indicating the general trend in the data; the vertical line in each panel is at the constituent boundary; time is normalised

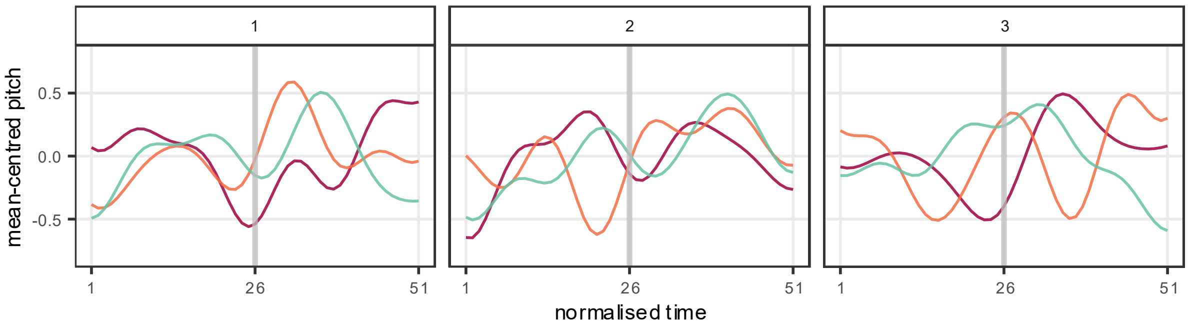


Figure 3: Pitch contours of three samples of three tokens each from the "right stress" cluster; each line represents one token; the vertical line in each panel is at the constituent boundary; time is normalised

## References:


**[1]** Bell, M. J., & Plag, I. (2012). Informativeness is a determinant of compound stress in English. *Journal of Linguistics, 48*(3), 485–520.
**[2]** Kunter, G. (2011). *Compound stress in English: The phonetics and phonology of prosodic prominence*. De Gruyter Mouton.
**[3]** Bell, M. J., & Plag, I. (2013). Informativity and analogy in English compound stress. *Word Structure, 6*(2), 129–155.
**[4]** Plag, I., Kunter, G., & Lappe, S. (2007). Testing hypotheses about compound stress assignment in English: A corpus-based investigation. *Corpus Linguistics and Linguistic Theory, 3*(2), 199–232.
**[5]** Plag, I., Kunter, G., Lappe, S., & Braun, M. (2008). The role of semantics, argument structure, and lexicalization in compound stress assignment in English. *Language, 84*(4), 760–794.
**[6]** Ostendorf, M., Price, P., & Shattuck-Hufnagel, S. (1996). *Boston University Radio Speech Corpus*. Linguistic Data Consortium.
**[7]** Plag, I., Kunter, G., & Schramm, M. (2011). Acoustic correlates of primary and secondary stress in North American English. *Journal of Phonetics, 39*(3), 362–374.2
**[8]** 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*, 4895891.
**[9]** Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, 4171–4186.