21. Phonetik und Phonologie im deutschsprachigen Raum

Dominic Schmitz[1], Melanie J. Bell[2] & Ingo Plag[1]

[1]Heinrich Heine University Düsseldorf, [2]Anglia Ruskin University Cambridge

{dominic.schmitz, ingo.plag}@uni-duesseldorf.de, melanie.bell@aru.ac.uk

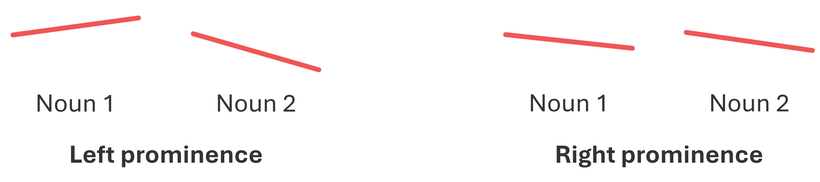
P&P 2025
PHONETIK UND PHONOLOGIE IM DEUTSCHSPRACHIGEN RAUM

English compound semantics are predictable by compound prominence and vice versa

Motivation

Received wisdom

- Compound stress
  - Left-stressed compounds: *cam**paign** promise* *__child__ care* *pro**ba**tion officer*
  - Right-stressed compounds: *Boston __har__bour* *home __phone__* *silk __tie__*
- About ⅓ of compounds are right-stressed [1-3]
- Many compounds are variable [4]
- Many factors influence stress perception [1-12]
  - semantic relation, specificity, and class, family size, analogy, lexicalization, length, speaker, …



Today's focus

- From categorical stress to pitch contours
- From categorical semantics to contextualized semantics, cf. [13] for Mandarin

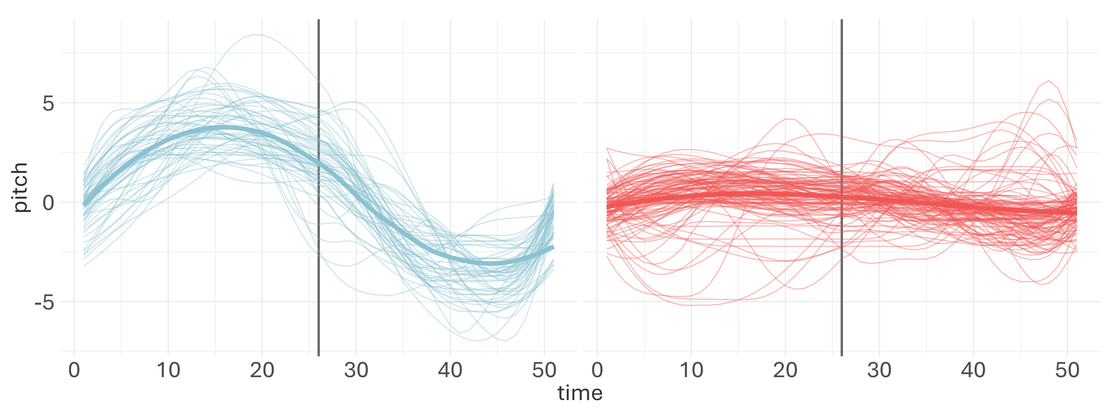
Data

- Boston University Radio Speech Corpus [14]
- American English, news texts, professional radio news presenters as speakers
- Noun-noun compounds: N = 4327, V = 2476
- Pitch data was extracted using sound files and TextGrids with the rPraat package [15]
- Processing steps: removal of octave jumps and erroneous pitch in voiceless segments, smoothing via the PraatSmooth algorithm of the soundgen package [16], transformation from Hz to semitones, centering by token, sampling at 51 equally spaced points in time

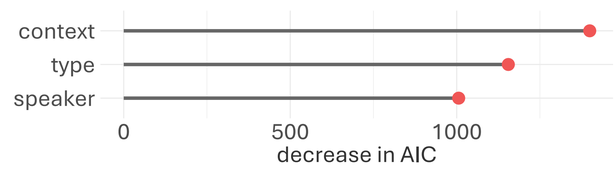
Sample

The device is attached to a plastic **wristband**. It looks like a watch. It functions like an electronic **probation officer**. When a computerized call is made to a former prisoner's **home phone**, that person answers by plugging in the device. The **wristband** can be removed only by breaking its clasp, and if that's done the inmate is immediately returned to jail. The description conjures up images of big brother watching. But Jay Ash, **deputy superintendent** of the Hampton County jail in Springfield, says the **surveillance system** is not that sinister.

Contours



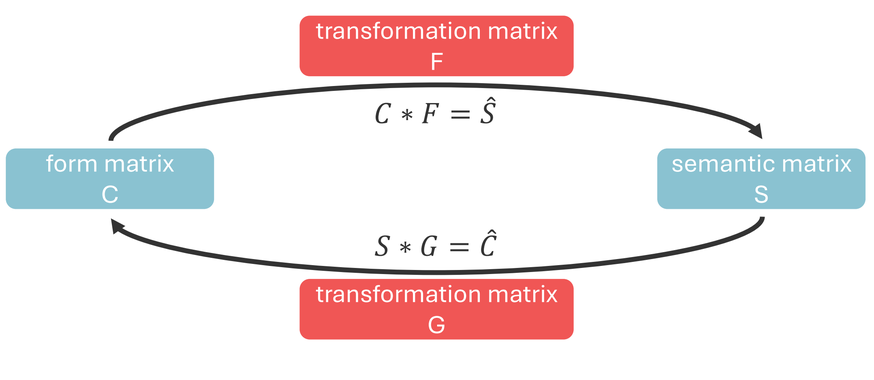
- Based on subset with N = 10, V = 167 to control for
  - context: each type occurs at least in 2 contexts, 27 contexts overall
  - speaker: each context is produced by at least 2 speakers
- Optimal clustering with k-means for longitudinal data: k = 2
- GAMs to model pitch as a function of time in interaction with predictors



Semantics

- Linear discriminative learning [17] to investigate whether pitch contours contain semantics
- Based on subset with N = 397, V = 79 to control for
  - context: each type occurs in at least 3 contexts
  - speaker: each context is produced by 1 speaker, at least 2 speakers per type
- How it works
  - Building associations between representations ('cues' and 'outcomes')
  - Association weight increases every time a given cue and a given outcome co-occur; decreases whenever the cue occurs without that outcome
  - Here: form = pitch data, meaning = context-dependent semantic vectors [18]
  - Train the network in leave-1-out style, predict the left-out token:

    Is the nearest neighbour of that token a token of the same type?



chance accuracy	**comprehension** form to semantics	**production** semantics to form
2 %	11.6 %	19.1 %

Discussion

- Pitch contours are predictable from contextually-informed semantics
- Semantics is predictable from pitch
- Predicting the pitch contours is easier than predicting the semantics
- Estimating 768 dimensions from 51 is more difficult and error-prone than the reverse
- Relationship between pitch and compound prominence is more complex than previous analyses suggest
- Speaker, compound, and context contribute to pitch contours
- Possibility of direct mapping from speech signal to experience-based semantics and vice-versa suggests language could involve a direct mapping between signal and semantics, with abstract categories being emergent

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]** Plag, I. (2010). Compound stress assignment by analogy: the constituent family bias. *Zeitschrift für Sprachwissenschaft, 29*(2), 243–282. **[4]** Bell, M. J. (2015). Inter-speaker variation in compound prominence. *Lingue e Linguaggio, XIV*(1/2015), 61–78. **[5]** Arndt-Lappe, S. (2011). Towards an exemplar-based model of stress in English noun–noun compounds. *Journal of Linguistics, 47*(3), 549–585. **[6]** Bell, M. J. (2015). Basic relations and stereotype relations in the semantics of compound nouns. *Journal of Cognitive Science, 16*(3), 224–260. **[7]** Bell, M. J., & Plag, I. (2013). Informativity and analogy in English compound stress. *Word Structure, 6*(2), 129–155. **[8]** Kunter, G., & Plag, I. (2007). What is compound stress? In *Proceedings of the 16th international congress of phonetic sciences* (pp. 1005–1008). **[9]** Plag, I. (2006). The variability of compound stress in English: Structural, semantic, and analogical factors. *English Language & Linguistics, 10*(1), 143–172. **[10]** 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. **[11]** 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. **[12]** Plag, I., & Kunter, G. (2010). Constituent family size and compound stress assignment in English. *Linguistische Berichte, Sonderheft* (17), 349–382. **[13]** Chuang, Y.-Y., Bell, M. J., Tseng, Y.-H., & Baayen, R. H. (2025). Word-specific tonal realizations in Mandarin. arXiv. **[14]** Ostendorf, M., Price, P., & Shattuck-Hufnagel, S. (1996). *Boston University Radio Speech Corpus*. Linguistic Data Consortium. **[15]** Bořil, T., Skarnitzl, R. (2016). Tools rPraat and mPraat. In *Text, speech, and dialogue*. Springer. **[16]** Anikin, A. (2019). Soundgen: An open-source tool for synthesizing nonverbal vocalizations. *Behavior Research Methods, 51*(2), 778-792. **[17]** 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. **[18]** Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 1, 4171–4186.