Heinrich Heine Universität Düsseldorf

Semantics and pitch contour predict one another in English compounds

Dominic Schmitz*, Melanie Bell#, & Ingo Plag*

*Heinrich-Heine-Universität Düsseldorf, #Anglia Ruskin University Cambridge

4th International Conference Prominence in Language, April 23-24, 2026, Cologne

A problem of form

Variable compound stress

ópera *glasses*	*Boston* ***márathon***
Óxford *Street*	*Oxford* ***Róad***
campáign *promise*	*summer* ***níght***

- About one third of the compounds are right-stressed

  (e.g. Bell & Plag 2012; Kunter 2011; Plag 2010)

- Many compounds are variable (Bell 2015)

  ice ***créam*** *-* ***íce*** *cream*

Factors influencing stress pattern

- Semantic relation between constituents
- Semantic specificity of constituents
- Semantic class of constituents
- Constituent family sizes
- Analogy
- Lexicalization
- Context
- Length
- Region
- Individual speaker
- ....

Arndt-Lappe 2011; Bell 2015a,b; Bell & Plag 2012, 2013; Kunter 2011; Kunter & Plag 2007; Plag 2006; Plag et al. 2007; Plag et al. 2008; Plag & Kunter 2010; Plag 2010

Factors influencing stress pattern

- Semantic relation between constituents
- Semantic specificity of constituents
- Semantic class of constituents
- Constituent family sizes
- Analogy
- Lexicalization
- Context
- Length
- Region
- Individual speaker
- ....

Explanations?

- How do speakers actually make use of these factors?
- How does the speaker learn (and know how) to apply these factors?

Hypothesis

- The observed effects emerge from a language system that originates in the speaker's experience

How to model experience?

- Discriminative learning

Discriminative Learning (Rescorla 1988 et seq.)

- Established learning theory, recently extended to language
  (Arnon & Ramscar 2012; Baayen et al. 2011, 2013, 2015; Baayen & Ramscar 2015; Blevins et al. 2015; Ramscar et al. 2010, 2013; Plag 2018)
- Learning results from exposure to informative relations among events in the environment (co-occurrence of cues and outcomes)
- Association weights, adjusted according to new, informative experiences (‘Rescorla-Wagner equations‘)
- Associations are strengthend upon the **presence** of a particular cue and the **presence** of a particular outcome
- Links are weakened upon the **presence** of a particular cue and the **non-presence** of a particular outcome (‘error-driven’)
- Discriminative Lexicon Model (Baayen and colleagues, 2019 et seq.)

Today's central question

How are aspects of form related to meaning?

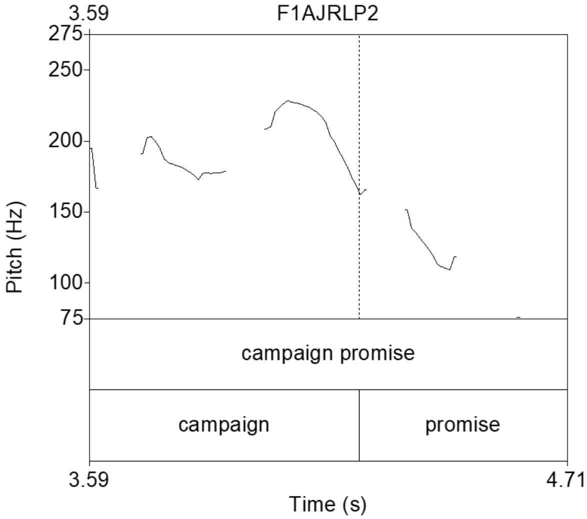
Pitch

- Pitch is the main correlate of compound stress in English
- Pitch contours may be directly related to (lexical) semantics
  (cf. Chuang et al. 2026; Li et al. 2026 for Mandarin)

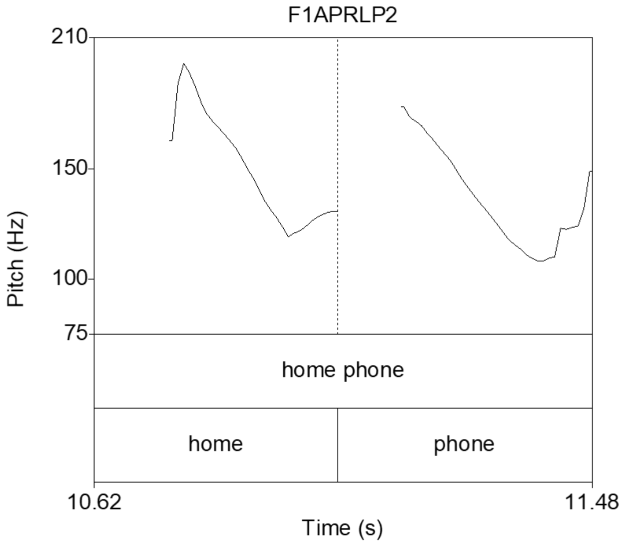
Roadmap

- Speech corpus data
- Pitch contours and compound stress
- Compound meaning: embeddings
- Relating form and meaning: Linear Discriminative Learning (LDL)

What about pitch?



one accent, on the left constituent



two accents, one on each constituent

Our study: Data

- Boston University Radio Speech Corpus (Ostendorf et al. 1996)
- American English, news texts, professional speakers
- NN compounds, N = 4327, V = 2476
- Used in previous studies of compound stress
  (e.g. Arndt-Lappe 2011; Bell & Plag 2012, 2013; Kunter 2011; Kunter & Plag 2007, Plag et al. 2008; Plag & Kunter 2010; Plag 2010)

Sample from BURSC

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.

Pitch data

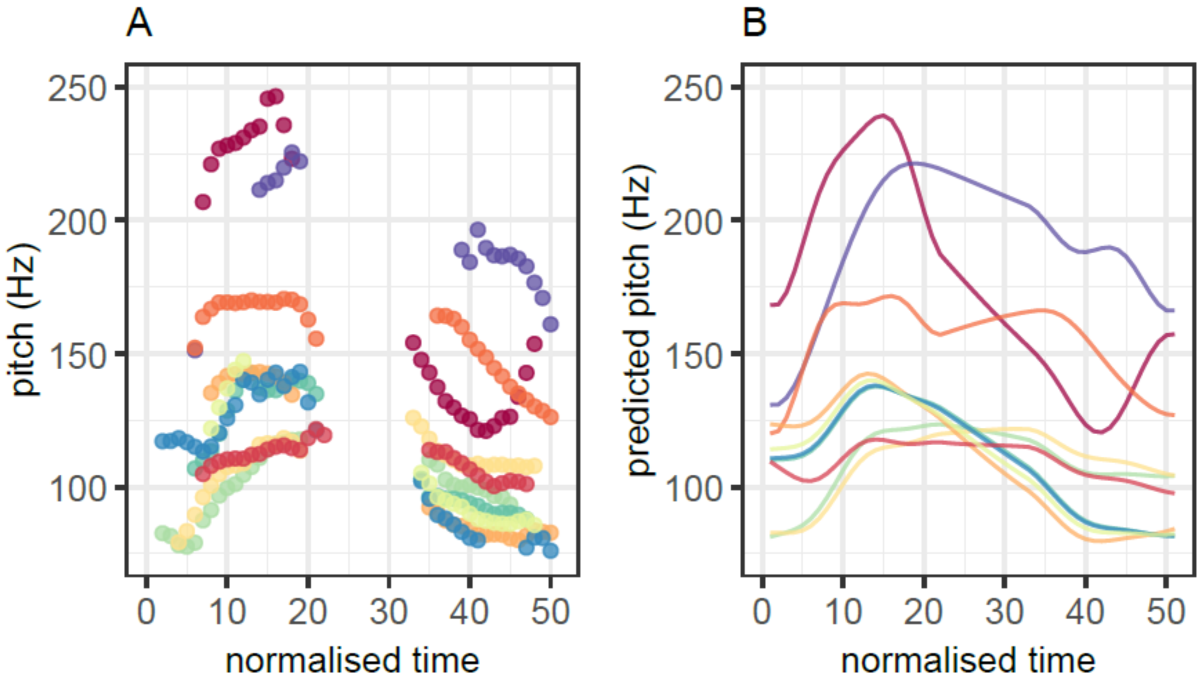
- Pitch was extracted using sound files and TextGrids, *rPraat* package in R (Bořil & Skarnitzl 2016)
- From raw to final **pitch** data
  - Sound to pitch with speaker-specific parameters
  - Removal of unvoiced frames
  - Time-normalization via 51 equally spaced points in time (across whole compound)
  - Time-normalization by constituent (26 points per constituent, averaging the middle one)
  - N=3,689, V=2,017

Arriving at pitch contours

- Generalised additive models (GAMs, Wood 2017)
- Can model complex wiggly curves, like pitch contours (Kösling et al. 2013; Chuang et al. 2024; Jin et al. 2025)
- Model pitch as function of time in interaction with various predictors
- Statistically informed estimates of the underlying, continuous pitch movements associated with each compound token

```
pitch ~   s(timeNorm, type) +
          s(timeNorm, speaker) +
          ti(timeNorm, durationN1) +
          ti(timeNorm, durationN2) +
          first order autoregression model('AR(1)')
```

Arriving at pitch contours



health care

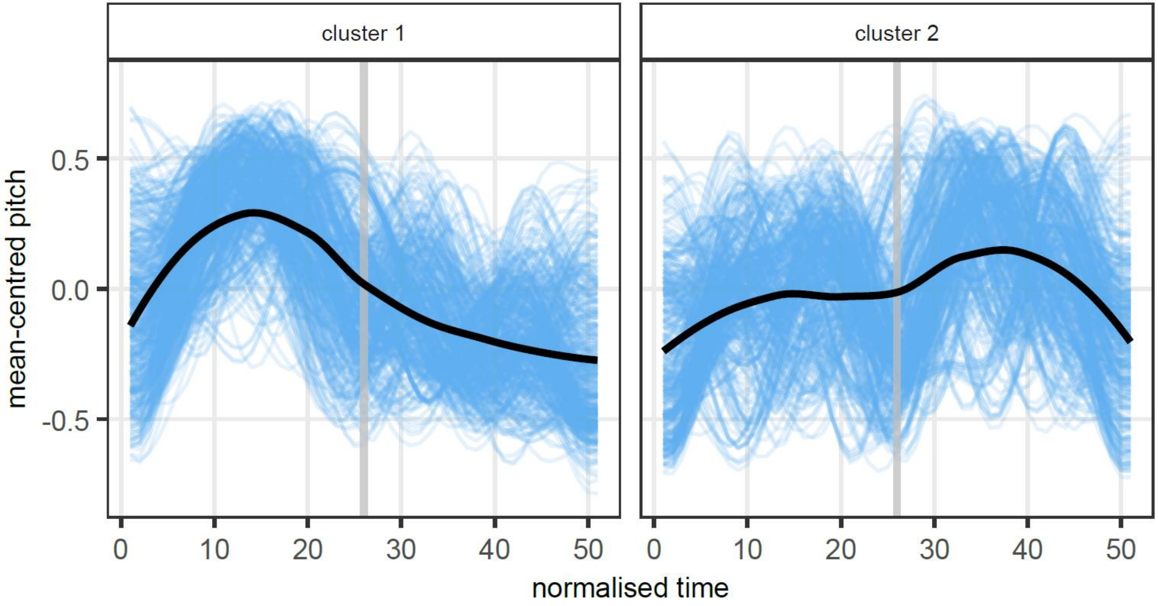
Arriving at a more controlled data set

Sampling from the corpus

- Types that occur in at least three different contexts
- Spoken by at least two speakers
- N=971, V=150 types

Two stress patterns? Pitch contours

- Optimal clustering with ‘k-means for longitudinal data’
- Optimal number of clusters: 2



Semantics and stress

- Semantic properties and semantic relations are predictive of stress patterns
  - ‘N1 refers to a period or point in time’
  - ‘N2 is a geographical term’
  - ‘N2 is a located at N1’
  - ‘N2 is made of N1’
  - ...
- Weak tendencies
- Vague definitions
- Ambiguity
- Theoretically questionable abstractions
- Coding of 18 semantic relations and eight semantic categories
- Six are predictive according to Plag et al. (2008)

Semantics: Embeddings

- Context-dependent token-level embeddings were extracted from BERT (Devlin et al. 2018)
- A transformers model pretrained on a very large corpus of English data in a self-supervised fashion
- Trained to guess the next word in sentences
- Which vector?
  Average, addition, multiplication, N2 embedding
- Outputs 768-dimensional vectors

Embeddings: Semantic matrix S (768 dimensions)

		[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]	[,10]	
wristband	[1,]	90	710	803	515	331	284	221	588	672	95	...
probation officer	[2,]	556	214	969	577	39	193	198	350	667	863	...
home phone	[3,]	844	62	194	157	894	186	496	497	723	614	...
	[...]	...										

Embeddings and stress patterns

Can the embeddings predict the stress pattern?

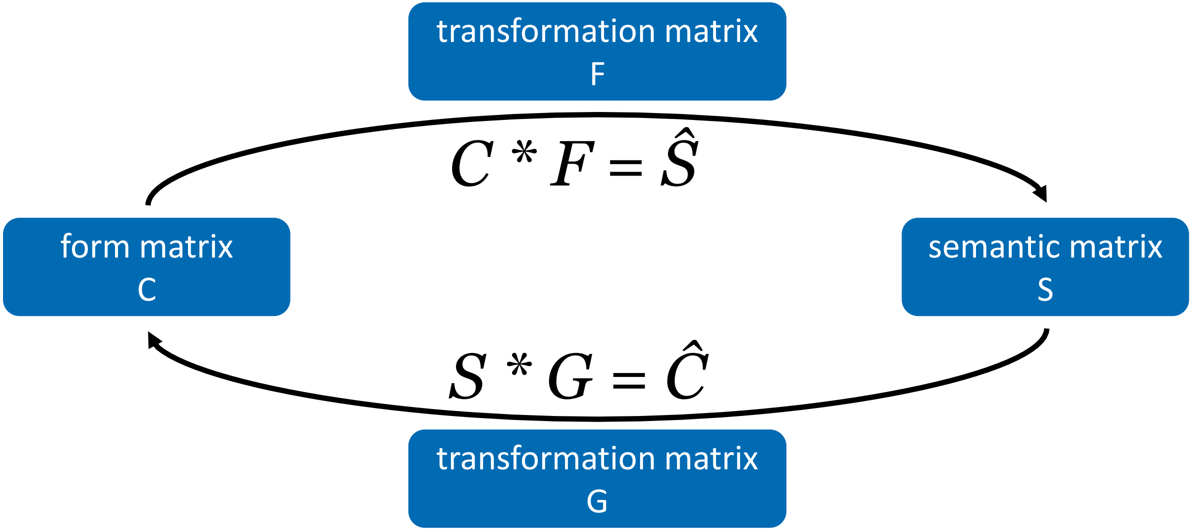
- Reduced data set (N=971, V=150)
- Logistic regression model to predict cluster membership directly from the semantic vectors

dimensions	C
768	1.00
88	0.77
50	0.75

It seems to be possible, with moderate success, to predict the stress pattern of compounds from their semantic vectors.

Mapping form on semantics, semantics on form

- Linear discriminative learning
  (Baayen et al. 2019; Julia packages: JudiLing, Luo et al. 2024; JudiLingMeasures, Heitmeier 2022)



The matrices

$$\mathbf{C} = \begin{array}{r} \\ \text{chief justice} \\ \text{retirement age} \end{array} \begin{array}{c} \begin{array}{ccccc} \texttt{cue1} & \texttt{cue2} & \texttt{cue3} & \texttt{cue4} & \texttt{cue5} \end{array} \\ \begin{pmatrix} 1.07 & 1.96 & 2.79 & 3.51 & 3.69 \\ 2.13 & 1.91 & 1.79 & 1.41 & 1.12 \end{pmatrix} \end{array}$$

Figure 7: Toy **C** matrix with acoustic cues.

$$\mathbf{S} = \begin{array}{r} \\ \text{chief justice} \\ \text{retirement age} \end{array} \begin{array}{c} \begin{array}{cccc} \texttt{S1} & \texttt{S2} & \texttt{S3} & \texttt{S4} \end{array} \\ \begin{pmatrix} 0.69 & 0.18 & 0.73 & 0.61 \\ 0.50 & 0.11 & 0.09 & 0.81 \end{pmatrix} \end{array}$$

Figure 8: Toy **S** matrix with semantic vectors

Predicting forms and predicting meanings

- 971 compound tokens in a leave-1-out design
- Train the transformation matrix on 970 compounds and then use the transformation matrix to predict the form (or semantics) of the 971st compound.
- 971 iterations, check accuracy of predictions
- ‘correct’: the nearest neighbour vector is that of a token of the same type

$\overrightarrow{deputy\ superintenent_1}$

$\overrightarrow{chief\ justice_2}$ $\overrightarrow{house\ speaker_3}$ — nearest neighbor

$\overrightarrow{chief\ justice_1}$ — nearest neighbor

Results: Accuracies

time normalization	**chance** accuracy	**comprehension** form to semantics	**production** semantics to form
across constituents	1.2 %		
within constituents	1.2 %		

Results: Accuracies

time normalization	**chance** accuracy	**comprehension** form to semantics	**production** semantics to form
across constituents	1.2 %	19.6 %	
within constituents	1.2 %	20.2 %	

Results

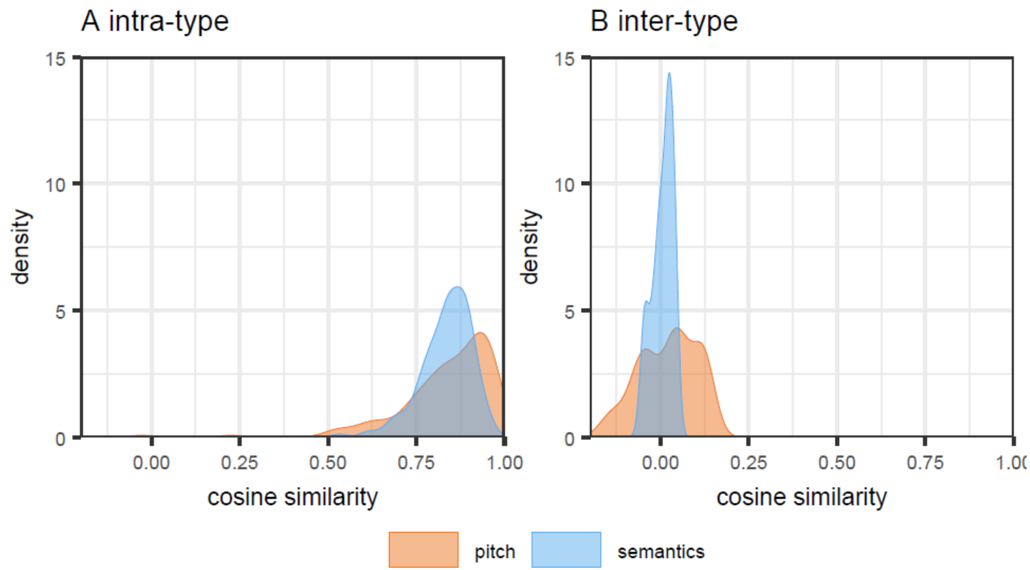
time normalization	**chance** accuracy	**comprehension** form to semantics	**production** semantics to form
across constituents	1.2 %	19.6 %	13.3 %
within constituents	1.2 %	20.2 %	12.5 %

Why does comprehension outperform production?

Understanding the results: The two spaces

- If tokens of a given type are more similar to each other, and
- types differ more clearly from other types in one of the two spaces,
- it is easier to detect the correct type in that space than in the other space.

Understanding the results: The two spaces



- Comprehension benefits from a semantic space that is both coherent within types and well separated across types.
- Production: the mapping of well-structured semantic vectors onto a much less structured pitch space

Summary and conclusion

- It is possible to predict the stress pattern from semantic embeddings
- It is possible to directly map from the speech signal to experience-based semantics, and from the semantics to the speech signal.
- There is indeed a close relationship between form (i.e. pitch) and lexical-contextual meaning in compounds.
- Compound stress is an emergent property from that mapping, based on the experience of the speakers.
- Previously found ‘predictors’ of compound stress may be epiphenomenal of this experience-based mapping.
- The gradience of the influence of the semantic effects is an inevitable consequence of the nature of the mapping.
- Compound stress is lexically learned, and this learning can be modelled.

Thank you

Funding:

Deutsche Forschungsgemeinschaft
Project ‘Form and meaning in English compounds‘, PL 151/13-1

References

Arndt-Lappe, S. (2011). Towards an exemplar-based model of stress in English noun–noun compounds1. *Journal of Linguistics*, *47*(3), 549–585. https://doi.org/10.1017/S0022226711000028

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. https://doi.org/10.1155/2019/4895891

Bell, M. J. (2015a). Basic Relations and Stereotype Relations in the Semantics of Compound Nouns. *Journal of Cognitive Science*, *16*(3), 224–260. https://doi.org/10.17791/JCS.2015.16.3.224

Bell, M. J. (2015b). Inter-speaker variation in compound prominence. *Lingue e Linguaggio*, *XIV*(1/2015), 61–78. https://doi.org/10.1418/80754

Bell, M. J., & Plag, I. (2012). Informativeness is a determinant of compound stress in English. *Journal of Linguistics*, *48*(3), 485–520. https://doi.org/10.1017/S0022226712000199

Bell, M. J., & Plag, I. (2013). Informativity and analogy in English compound stress. *Word Structure*, *6*(2), 129–155. https://doi.org/10.3366/word.2013.0042

Bořil, T., & Skarnitzl, R. (2016). *Tools rPraat and mPraat* (pp. 367–374). https://doi.org/10.1007/978-3-319-45510-5_42

Heitmeier, M. (2022). *JudiLingMeasures*. Eberhard Karls Universität, Seminar für Sprachwissenschaft.

Kunter, G. (2011). *Compound stress in English: The phonetics and phonology of prosodic prominence*. De Gruyter Mouton.

Kunter, G., & Plag, I. (2007). What is compound stress? In J. Trouvain & J. Barry (Eds.), *Proceedings of the 16th international congress of phonetic sciences* (pp. 1005–1008).

Luo, X., Heitmeier, M., Chuang, Y.-Y., & Baayen, R. H. (2024). *JudiLing: an implementation of the Discriminative Lexicon Model in Julia*. Eberhard Karls Universität, Seminar für Sprachwissenschaft.

Ostendorf, M., Price, P., & Shattuck-Hufnagel, S. (1996). *Boston University Radio Speech Corpus*. Linguistic Data Consortium.

Plag, I. (2006). The variability of compound stress in English: structural, semantic, and analogical factors. *English Language & Linguistics*, *10*(1), 143–172. https://doi.org/10.1017/S1360674306001821

Plag, I. (2010). Compound stress assignment by analogy: the constituent family bias. *Zeitschrift Für Sprachwissenschaft*, *29*(2), 243–282.

Plag, I., & Kunter, G. (2010). Constituent family size and compound stress assignment in English. *Linguistische Berichte*, *Sonderheft*(17), 349–382.

Plag, I., Kunter, G., & Lappe, S. (2007). Testing Hypotheses About Compoundstress Assignment in English: A Corpus-Based Investigation. *Corpus Linguistics and Linguistic Theory*, *3*(2), 199–232. https://doi.org/10.1515/CLLT.2007.012/MACHINEREADABLECITATION/RIS

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. https://doi.org/10.1353/LAN.0.0072

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). *Language Models are Unsupervised Multitask Learners*.