SFB 1102: C1
Information Density and the Predictability of Phonetic Structure

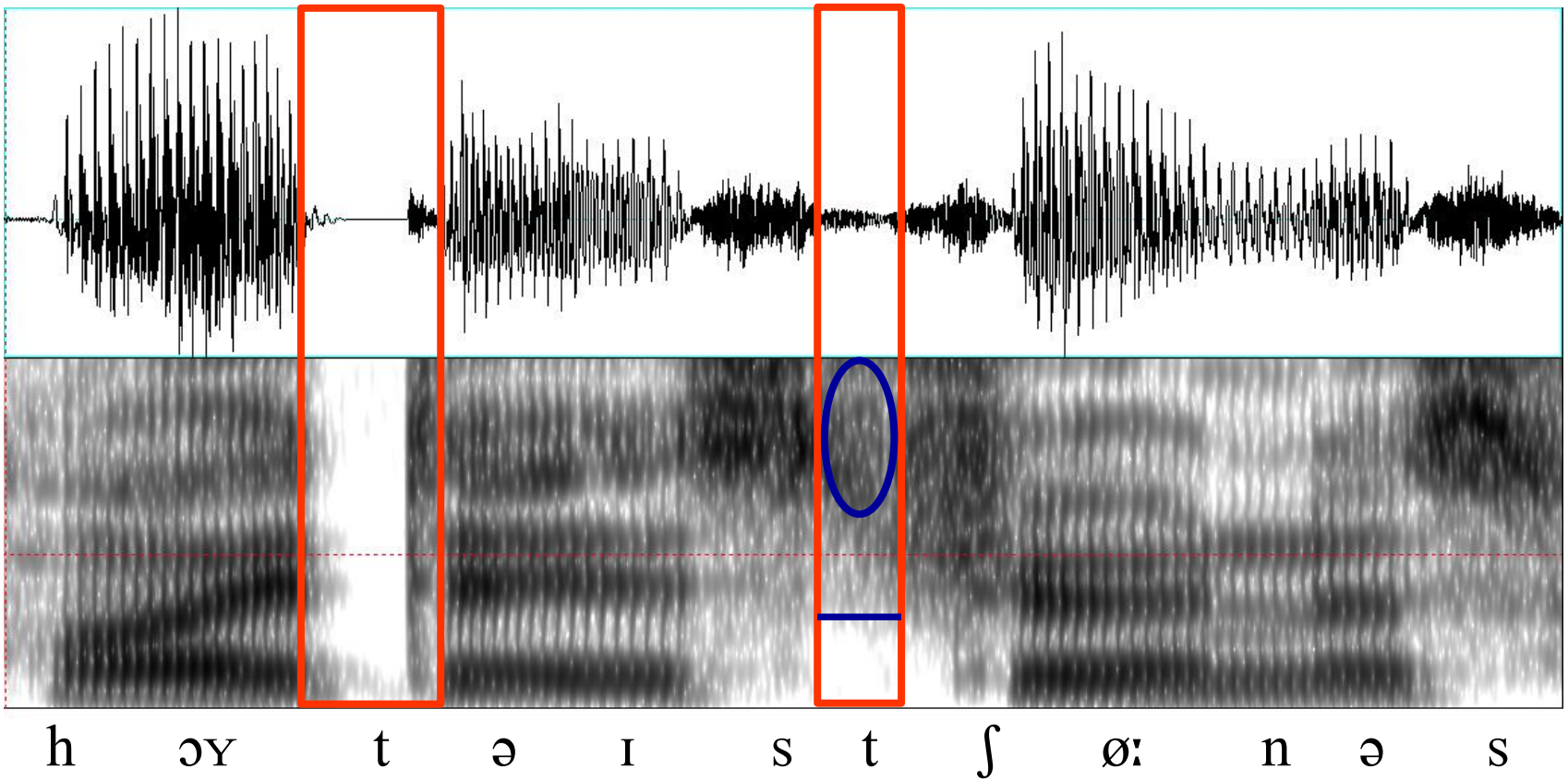
E. Schulz, Y.M. Oh, Z. Malisz, B. Andreeva, and B. Möbius (2016). Impact of prosodic structure and information density on vowel space size. Proc. Speech Prosody 2016, 350-354.

Z. Malisz, E. Brandt, B. Möbius, Y.M. Oh, and B. Andreeva (2018). Dimensions of Segmental Variability: Interaction of Prosody and Surprisal in Six Languages, Frontiers in Communication / Language Sciences, vol. 3, article 25, 1-18, doi.org/10.3389/fcomm.2018.00025.

IDeaL SFB 1102

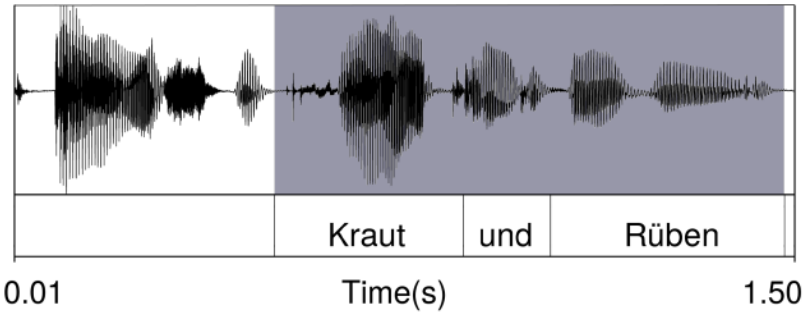
UNIVERSITÄT DES SAARLANDES

Example: [t] illusion

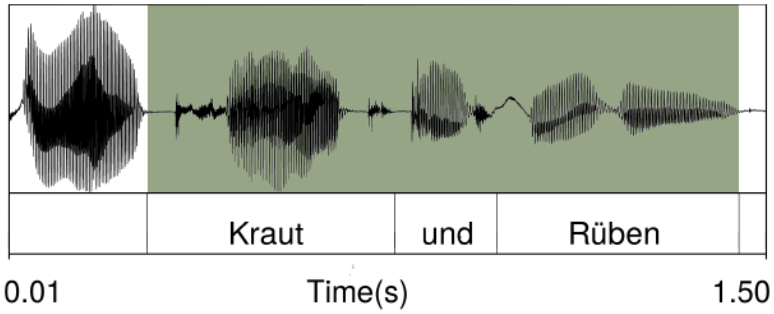


Example: predictability by collocation

collocation

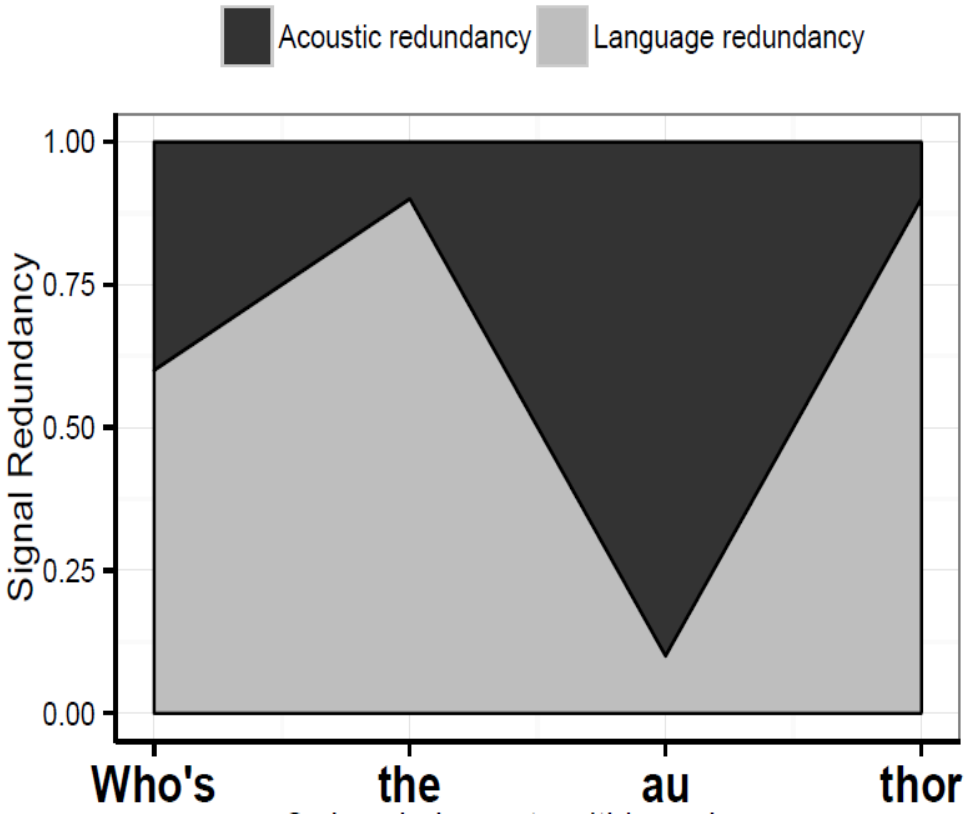


item list



Kraut und Rüben used as **idiom**: **Temporal reduction**.

Smooth Signal Redundancy Hypothesis (SSRH) (Aylett & Turk 2004)

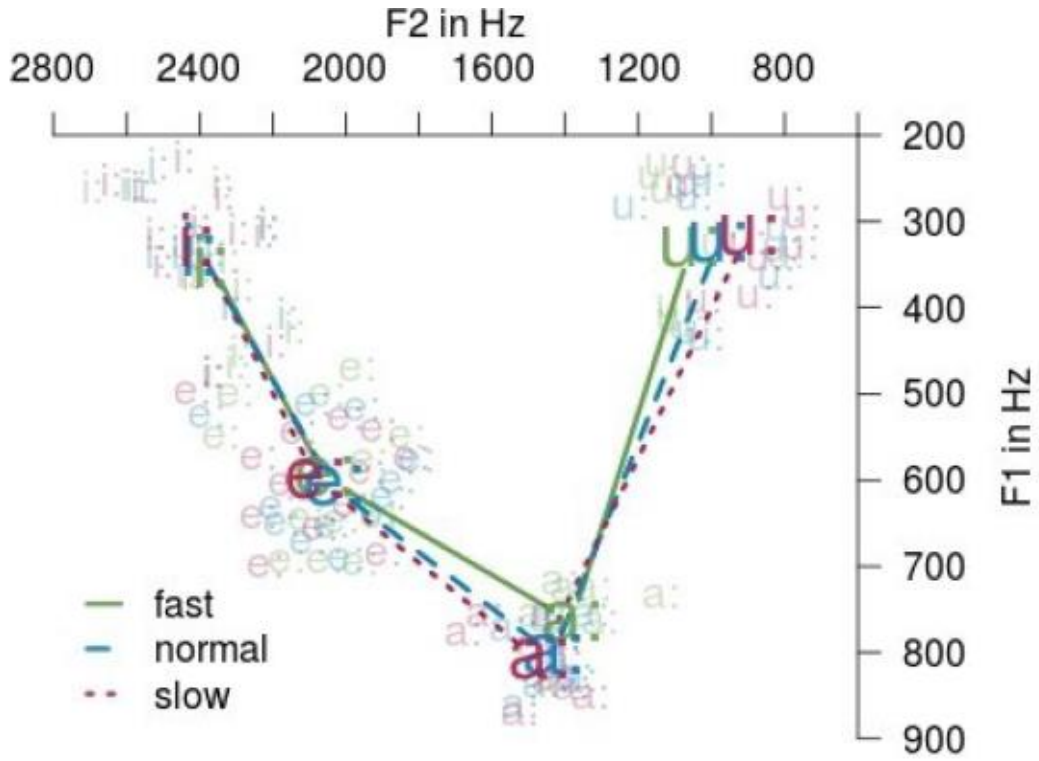


Negative relationship between language redundancy and acoustic redundancy.

Result: Communication becomes more robust by distributing information more smoothly over the signal.

Vowel space size

- Sex
  (Fant 1966)
- Speaking style
  (Bradlow, Kraus, and Hayes 2003)
- Language redundancy
  (Aylett and Turk 2006)
- Prosodic structure
  (Bergem 1993)
- Speech rate
  - Reduced size with increased speech rate
    (Turner, Tjaden, and Weismer 1995; Weiss 2007)
  - Minimal or no effect
    (Son and Pols 1990; Fourakis 1991)



German female vowel space at slow, normal and fast speech rate.

Corpus study: Aim

Aim

1. Influence of information density and prosodic structure on vowel space size
2. Influence of speech rate on vowel space size
3. Investigate these phenomena from cross-linguistic perspective

Corpus study: Method

Material

- BonnTempo Corpus (Dellwo et al. 2004)
- FRA, DEU, FIN, CES, POL (3 females, 3 males)
- Intended speech rates: slow, normal, fast

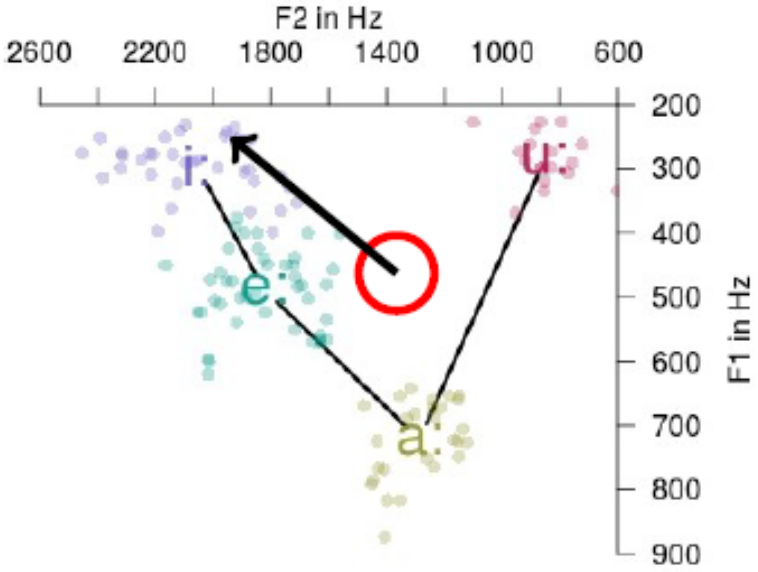
German original

Am nächsten Tag fuhr ich nach Husum. Es ist eine Fahrt ans Ende der Welt; hinter Gießen werden die Berge und Wälder eintönig, hinter Kassel die Städte ärmlich, und bei Salzgitter wird das Land flach und öde. Wenn bei uns Dissidenten verbannt würden, würden sie ans Steinhuder Meer verbannt.

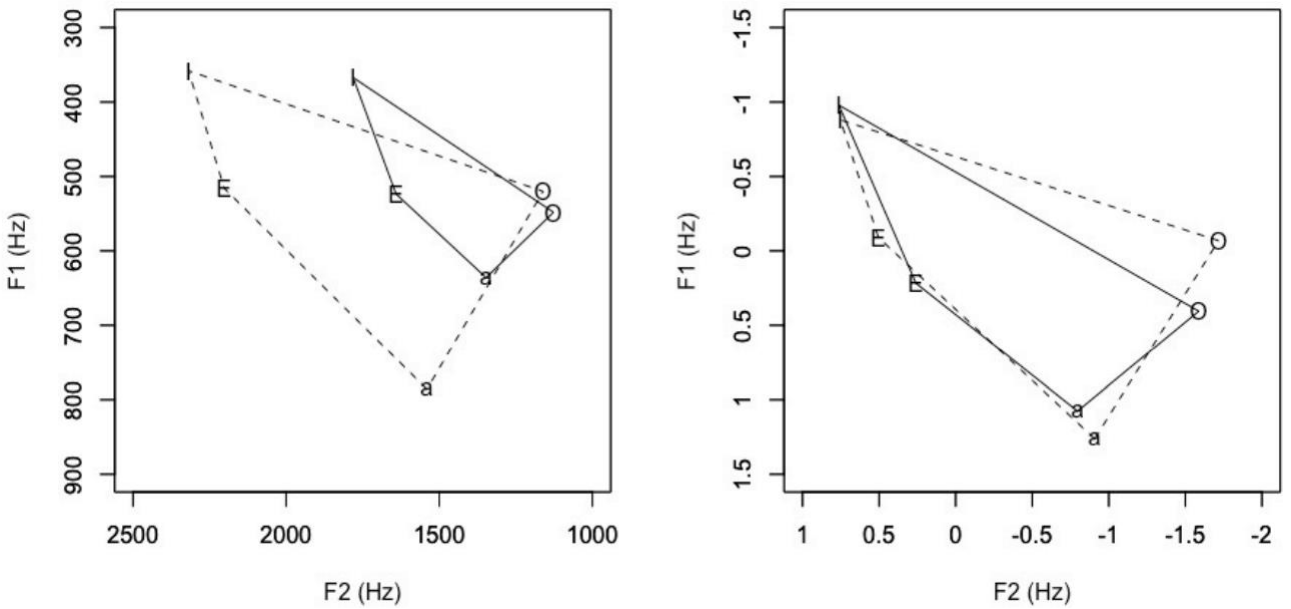
Bernhard Schlink (1994) *Selbs Betrug*: Diogenes (p. 242)

Corpus study: Spectral analysis

- F1 and F2 measured at temporal midpoint of vowel with Praat (Boersma and Weenink 2015)
- Speaker-dependent normalization of formant values (Lobanov 1971)
- Vowel dispersion = Euclidean distance between midpoint of the vowel space and formant values for every vowel which is calculated for each speaker (Bradlow, Torretta, and Pisoni 1996)
- Following other work on variability of vowel space size (Munson and Solomon 2004; Bergem 1993)



Lobanov Normalisierung



$$Fn.norm = (Fn - Fn.mean)/Fn.sd$$

Fn.mean, Fn.sd: Mittelwert und Standardabweichung von Fn über eine Stichprobe von Vokalen desselben Sprechers.

Lobanov, B. (1971). Classification of Russian vowels spoken by different speakers. *Journal of the Acoustical Society of America*, 49, 606–608.

Corpus study: Language redundancy model

- Biphone language models for each language were built

Language	Corpus	N tokens
CES	Frequency dictionary	398.0 mill.
DEU	Frankfurter Rundschau	52.0 mill.
ENG (AmE)	COCA	540.0 mill.
FIN	Finnish PAROLE	180.0 mill.
FRA	LEXIQUE 3.80	9.1 mill.
POL	Frequency dictionary	901.0 mill.

Corpus study: Language redundancy model

- Biphone language models for each language were built
- Surprisal was estimated taking the previous context into account $(X\_n|X\_n-1)$

Surprisal

- Measures the surprise of encountering a linguistic unit Xn in a specific context Xc
- Based on language models (LMs) which estimate the distribution of sequences of linguistic units in a language

$$Surprisal(Phone_i) = -log_2 P(Phone_i|Phone_{i-1})$$

Corpus study: Prosodic model

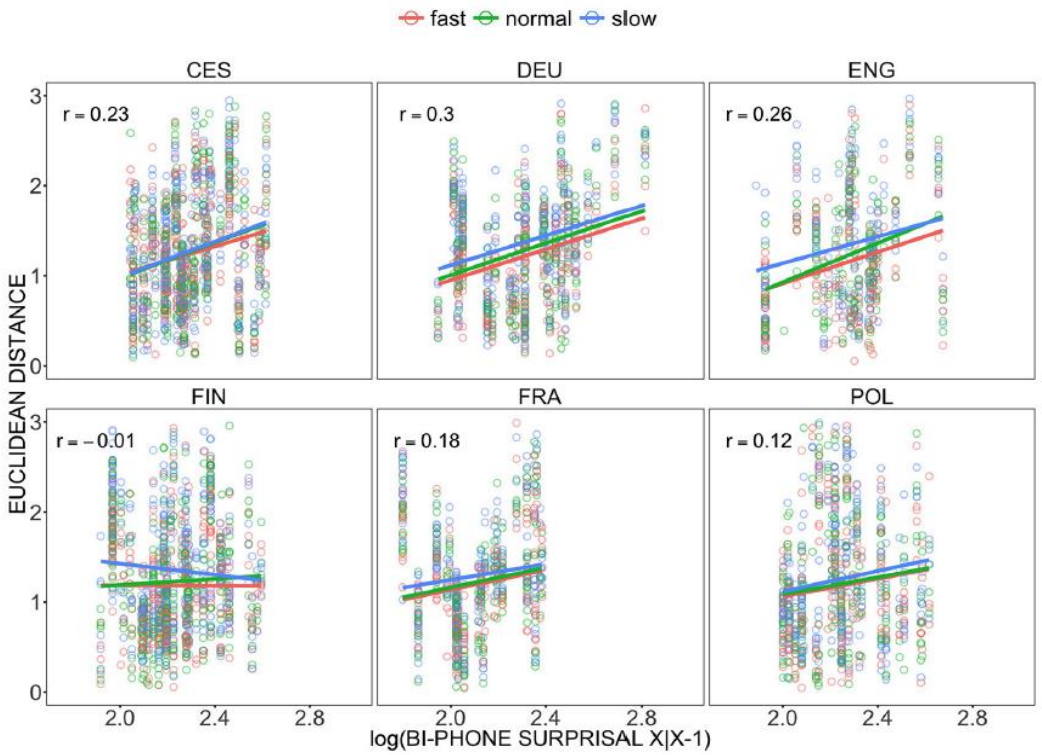
- largely similar to Aylett and Turk (2006)
  1. Prominence
     - none
     - primary lexical stress
     - if monosyllabic: function words unstressed, content words stressed
     - not included: high probability of having a phrasal accent
  2. Boundary
     - none
     - word boundary
     - high probability of phrase boundary (followed by pause of $\geq$ 100 ms)

Results: Correlation analysis

- Positive relationship between vowel dispersion and surprisal of biphone $X\_n|X\_n-1$

Pearson's correlation

$r = 0.17$, $p < .001$



Relationship between vowel dispersion and information density for each language.

Results: Linear mixed effects model

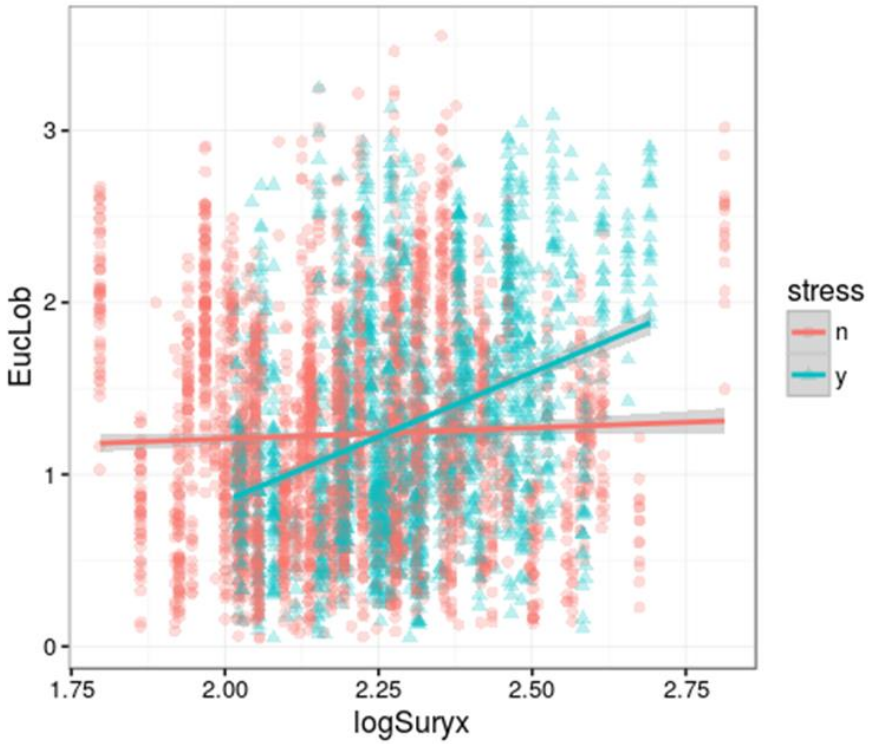
Lob(EucDist) ~ Surprisal * VowelID + SpeechRate + Boundary + Prominence + (1+Surprisal|Word) + (1|Language)

Stress (y-n)	0.04	0.03	1.28	=0.20
Boundary (y-n)	−0.04	0.04	−1.08	=0.28
Tempo (normal-fast)	0.04	0.01	3.20	=0.001
Tempo (slow-fast)	0.12	0.01	10.06	<0.001
Surprisal-preceding	0.70	0.27	2.65	=0.009
Vowel identity (/a/-Mean)	−0.11	0.03	−3.43	=0.001
Vowel identity (/e/-Mean)	−0.60	0.03	−17.18	<0.001
Vowel identity (/i/-Mean)	0.35	0.04	8.49	<0.001
Boundary*Surprisal	0.07	0.14	0.47	=0.64
Stress*Surprisal	0.21	0.24	0.89	=0.37
Tempo (normal)*Surprisal	0.08	0.06	1.24	=0.21
Tempo (slow)*Surprisal	0.01	0.06	0.20	=0.84

Results: Linear mixed effects model

Lob(EucDist) ~ Surprisal * VowelID + SpeechRate + Boundary + Prominence + (1+Surprisal|Word) + (1|Language)

1. Language redundancy
   - significant effect for surprisal
2. Prosodic structure
   - Vowel space expanded from fast to normal, and from normal to slow speech rate
   - Vowel dispersion increased under lexical stress
   - Vowel dispersion increased for vowels before phrase boundaries compared to no boundary



Discussion

- Averaged over all languages: Positive relationship between surprisal and vowel dispersion
- Does not hold for all languages
  - Polish: has only weak spectral vowel reduction (Nowak 2006)
  - Finnish: vowel reduction realised through duration (Suomi, Toivanen, and Ylitalo 2008)
- Support for literature that found effect of speech rate on vowel space size
- Prosodic model of prominence and boundary gives expected results

Summary

Effects of ID on phonetic encoding

- Small but robust, after controlling for basic prosodic factors (read speech)
- Robust across speech rates and languages

References

Aylett, M. and A. Turk. "Language redundancy predicts syllabic duration and the spectral characteristics of vocalic syllable nuclei". In: *Journal of the Acoustical Society of America* 119 (2006), pp. 3048–3058.

Bergem, D. R. van. "Acoustic vowel reduction as a function of sentence accent, word stress, and word class". In: *Speech Communication* 12 (1993), pp. 1–23.

Boersma, P. and D. Weenink. *Praat: doing phonetics by computer [Computer program]. Version 5.4.22*. 2015. URL: http://www.praat.org/.

Bradlow, A. R., N. Kraus, and E. Hayes. "Speaking clearly for children with learning disabilities: Sentence perception in noise". In: *Journal of Speech, Language, and Hearing Research* 46 (2003), pp. 80–97.

Bradlow, A. R., G. M. Torretta, and D. B. Pisoni. "Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics". In: *Speech Communication* 20.3 (1996), pp. 255–272.

Dellwo, V. et al. "BonnTempo-corpus and BonnTempo-tools: a database for the study of speech rhythm and rate". In: *Interspeech 2004*. 2004, pp. 777–780.

Fant, G. "A note on vocal tract size factors and non-uniform F-pattern scalings". In: *STL-QPSR* 4 (1966), pp. 22–30.

Fourakis, M. "Tempo, stress, and vowel reduction in American English". In: *Journal of the Acoustical Society of America* 90.4 (1991), pp. 1816–1827.

Lobanov, B. M. "Classification of Russian vowels spoken by different speakers". In: *Journal of the Acoustical Society of America* 49 (1971), pp. 606–608.

Munson, B. and N. P. Solomon. "The effect of phonological neighborhood density on vowel articulation". In: *Journal of Speech, Language, and Hearing Research* 47 (2004), pp. 1048–1058.

Nowak, P. M. "Vowel reduction in Polish". PhD thesis. University of California, Berkeley, 2006.

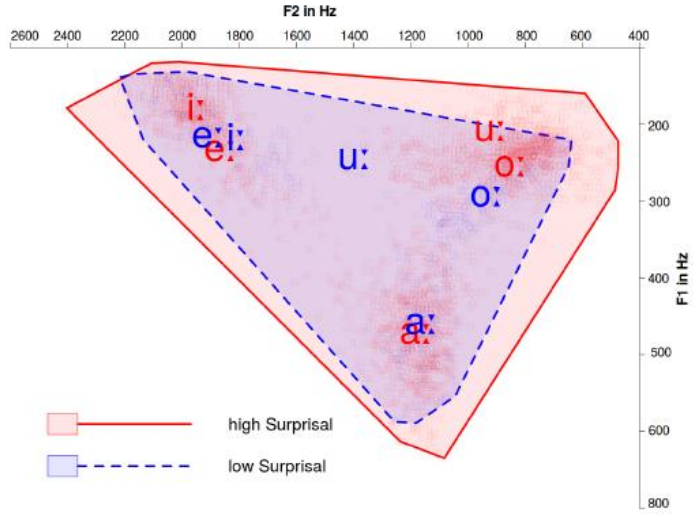
Son, R. J. J. H. van and L. C. W. Pols. "Formant frequencies of Dutch vowels in a text, read at normal and fast rate". In: *Journal of the Acoustical Society of America* 88 (1990), pp. 1683–1693.

Suomi, K., J. H. Toivanen, and R. Ylitalo. *Finnish sound structure. Phonetics, phonology, phonotactics and prosody*. Uolo University Press, 2008.

Turner, G. S., K. Tjaden, and G. Weismer. "The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis". In: *Journal of Speech and Hearing Research* 38 (1995), pp. 1001–1013.

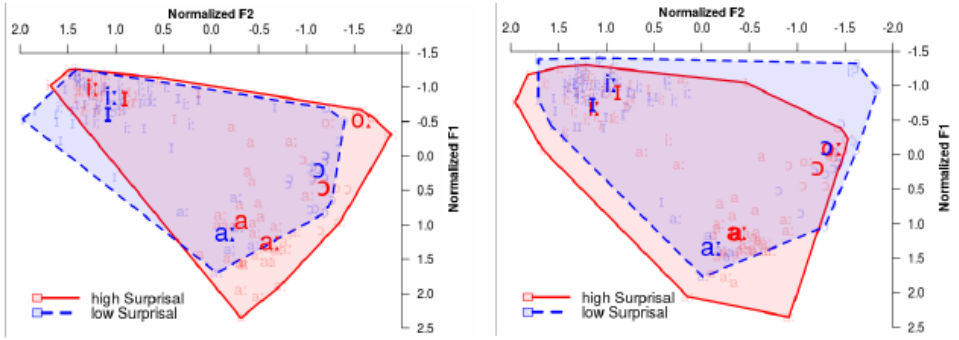
Weiss, B. "Rate Dependent Vowel Reduction in German". In: *Proceedings of the 12th SPECOM*. Moscow, 2007.

Vowel space: Distinctiveness



German **vowel space expands** in **high Surprisal** as opposed to **low Surprisal** conditions.

Vowel space of Bulgarian L2 speakers of German



- Advanced L2 speakers (left) expand the vowel space under **high Surprisal** conditions almost like German L1 speakers.
- Intermediate L2 speakers (right) apply native BG reduction patterns under **low Surprisal**.