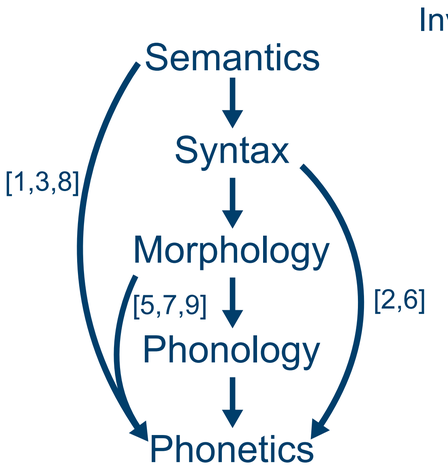
How meaning affects the duration of Japanese homophonous words

Motoki Saito[1] & Ruben van de Vijver[2]

[1]Carl von Ossietzky Universität Oldenburg
[2]Heinrich Heine Universität Düsseldorf

Background



Investigated predominantly ...

English
where
duration is
not contrastive

Japanese
where
duration is
contrastive

Aims

1. Does **homophone duration** covary with **semantics** also in **Japanese**?
2. Are **semantic** effects tied to **lexicality** of words?

Main Finding

Does semantics affect homophone duration in Japanese?

YES!

Data

Corpus of Spontaneous Japanese (CSJ) [10]

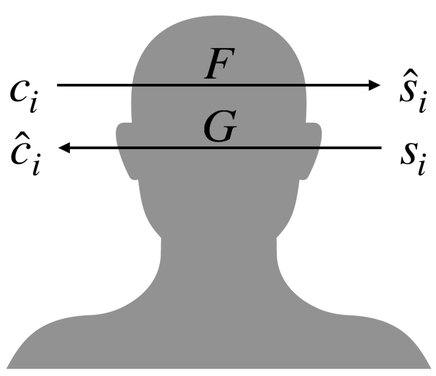
- The "core" section
- 44 hours of speech
- 500,000 words
- Mostly formal monologues

99,776 homophonous word tokens

- 1,586 word types in orthography
- 1,200 word types in phonetic transcriptions

Analysis

Discriminative Lexicon Model



(Unconditional) Semantic Support

<PROG> → /-ɪŋ/ → Less uncertainty
→ Greater semantic support

<PAST> → /-d/, /-t/, /-əd/, /-ɔːt/ → Greater uncertainty
→ Less semantic support

Conditional Semantic Support

e.g., goggles → -s is more predictable
→ Less conditional semantic support for -s

e.g.., suns → -s is less predictable
→ Greater conditional semantic support for -s

Generalized additive mixed-effects models

✓ Model 1: **Word**Dur ~ s(**u**SemSup) + Covariates
Model 2: **Word**Dur ~ s(**c**SemSup) + Covariates
Model 3: **Mora**Dur ~ s(**u**SemSup) + Covariates
✓ Model 4: **Mora**Dur ~ s(**c**SemSup) + Covariates

Covariates:
s(SpRate) + s(Freq) + s(BimoraFreq) + UttBgn
+ UttEnd + PoS + Gender + s(Speaker, bs='re')

SpRate: Numbers of moras / durations of utterances
Freq: Word frequency from CSJ
BimoraFreq: Sum of bimora frequency / word length
UttBgn & UttEnd: Utterance-initial/-final positions
PoS: Parts-of-speech
Speaker: Speaker (as a random intercept)

Semantic & form representations

/ʑiɕiɴ/	S001	S002	S003	S004	S005
自身 "self"	-0.34	0.78	0.61	0.45	0.09
自信 "confidence"	0.22	-0.37	-0.10	0.77	0.36
地震 "earthquake"	0.01	0.56	-0.86	-0.00	0.34

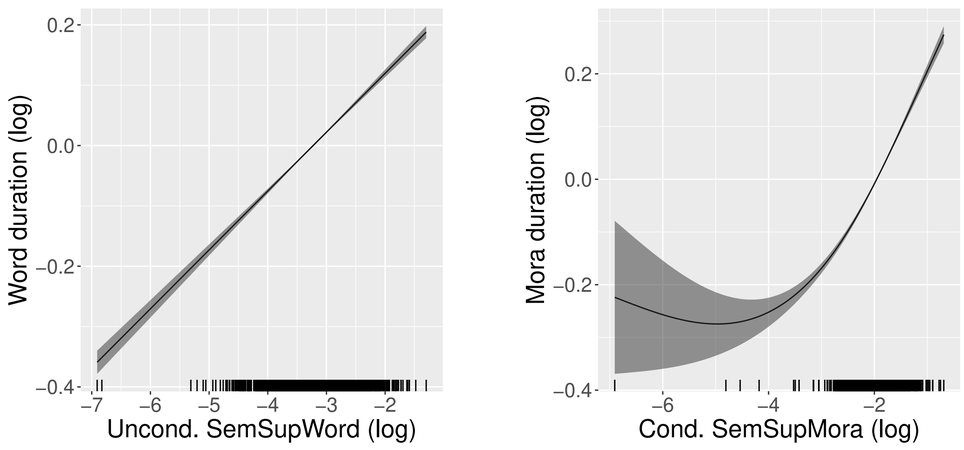
Semantic representation: A pre-trained fastText model [4]

Linear mapping

/ʑiɕiɴ/	#ji]shi #ジ]シ	ji]shin ジ]シン	shin# シン#	#jishi #ジシ	jishin ジシン
自身 "self"	1	1	1	0	0
自信 "confidence"	0	0	1	1	1
地震 "earthquake"	0	0	1	1	1

Form representation: Tri-moras with pitch accents

Results



Unconditional semantic support → better for **word** duration

Conditional semantic support → better for **mora** duration

Greater **unconditional** semantic support
→ Longer **word** duration

Greater **conditional** semantic support
→ Longer **mora** duration

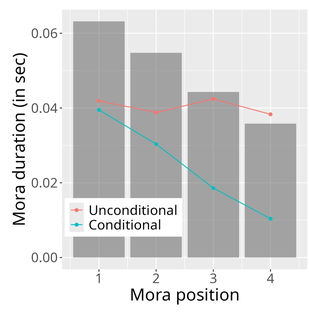
Discussion

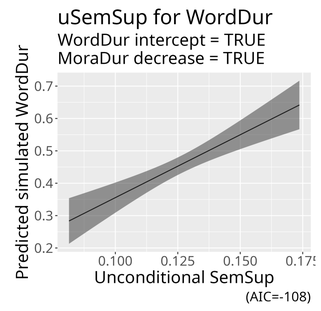
Homophone duration → **correlated with certainty** between semantics & forms.
also in a **mora-timed language** with **durational contrasts.**

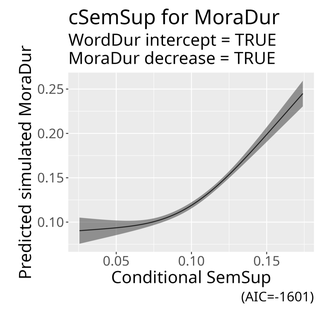
But, **why:** is **unconditional** semantic support better for **word** duration?
is **conditional** semantic support better for **mora** duration?

Unconditional semantic support → **Word-level idiosyncracy**
Conditional semantic support → **Decreasing mora duration** within a word

Similar u/cSemSup effects were observed only when
Simulations confirmed: 1) each **word** has its own **durational intercept** and
2) **mora duration decreases** within a word.







References

[1] 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, 1–39.
[2] Gahl, S. (2008). Time and thyme are not homophones: the effect of lemma frequency on word durations in spontaneous speech. Language, 84(3), 474–496.
[3] Gahl, S., & Baayen, R. H. (2024). Time and thyme again: Connecting english spoken word duration to models of the mental lexicon. Language, 100(4), 623–670.
[4] Grave, E., Bojanowski, P., Gupta, P., Joulin, A., & Mikolov, T. (2018). Learning word vectors for 157 languages. Proceedings of the International Conference on Language Resources and Evaluation (LREC 2018).
[5] Hay, J. (2007). The phonetics of 'un'. In J. Munat (Ed.), Lexical creativity, texts and contexts (pp. 39–57). John Benjamins.
[6] Lohmann, A. (2018). Cut (n) and cut (v) are not homophones: lemma frequency affects the duration of noun–verb conversion pairs. Journal of Linguistics, 54(4), 753–777.
[7] Plag, I., Homann, J., & Kunter, G. (2017). Homophony and morphology: The acoustics of word-final S in English. Journal of Linguistics, 53(1), 181–216.
[8] Saito, M., Tomaschek, F., Sun, C.-C., & Baayen, R. H. (2024). Articulatory effects of frequency modulated by semantics. In M. Schlechtweg (Ed.), Interfaces of phonetics (pp. 125–154). De Gruyter Mouton.
[9] Smith, R., Baker, R., & Hawkins, S. (2012). Phonetic detail that distinguishes prefixed from pseudo-prefixed words. Journal of Phonetics, 40(5), 689–705.
[10] The National Institute for Japanese Language. (2006). Construction of the corpus of spontaneous japanese (Vol. 124).