

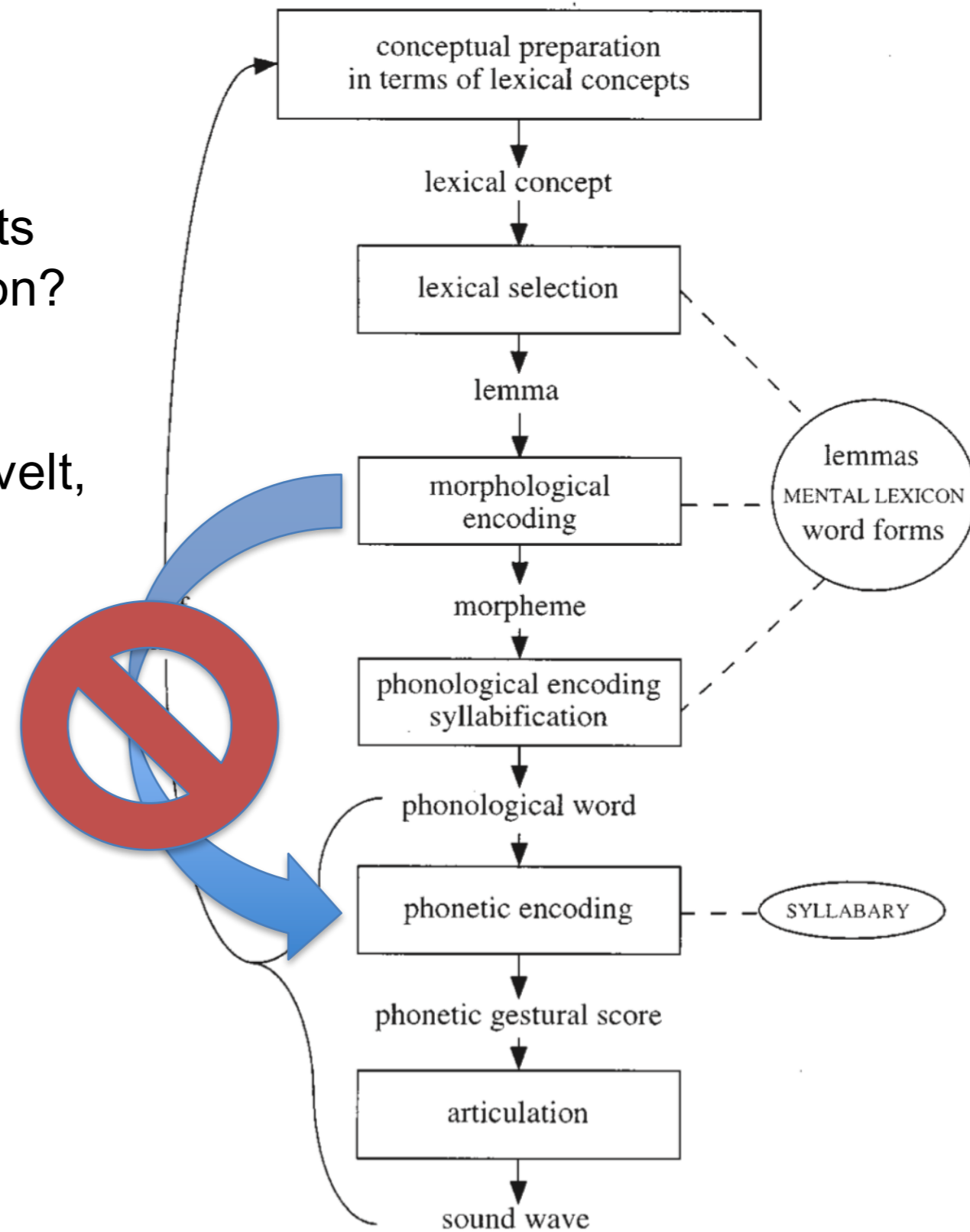
Morphological effects on the acoustics of Dutch /s/

Tim Zee

Background

Does the morphological status of segments influence their production and/or perception?

Traditional models of word production (Levelt, Roelofs & Meyer, 1999) predict that this should not be the case.



Background

- However, for American English word-final /s/ some evidence has been found that this might be the case (Plag, Homann & Kunter, 2017; Tomaschek et al., n.d.)

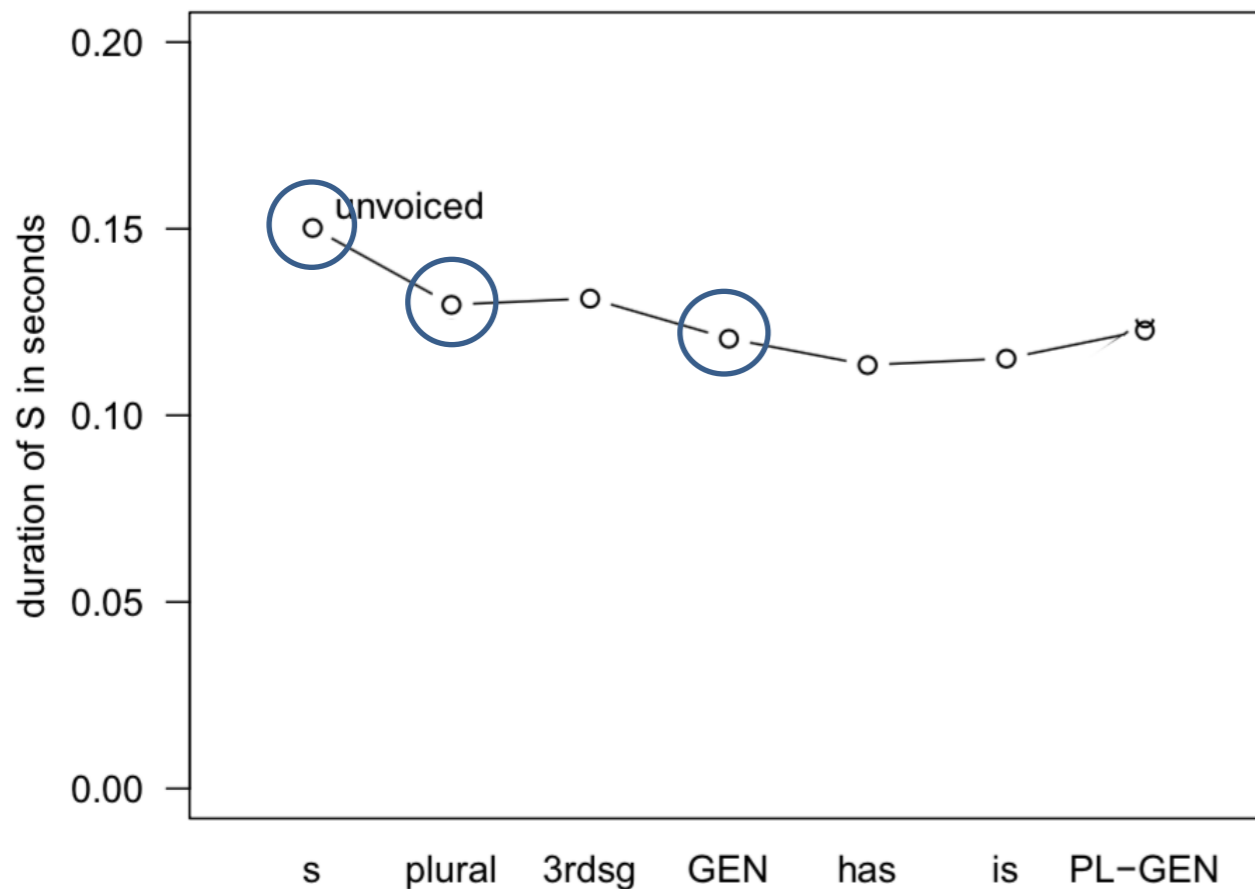


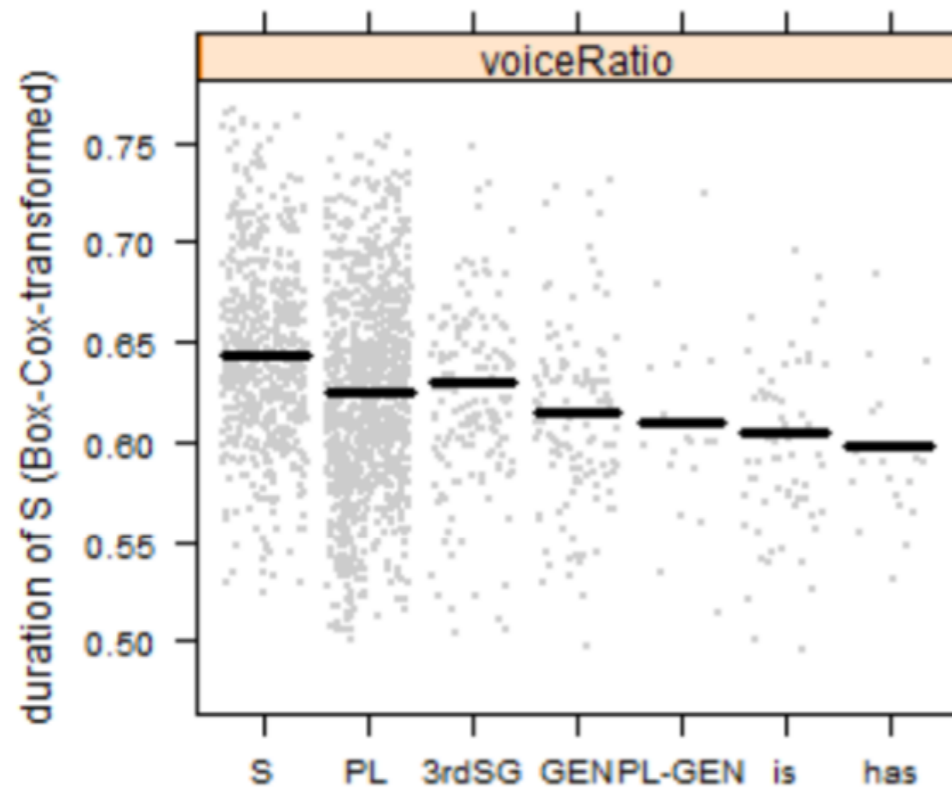
Figure 4

Interaction of type of S and voicing, Model 1 (Abbreviations: s = non-morphemic S, 3rdsg = 3rd person singular, GEN = genitive, PL-GEN = genitive-plural).

	S	PL	3RDSG	GEN	HAS/IS	PL-GEN
S		yes	yes	yes	yes	no
PL					yes	
3RDSG					yes	
GEN						
HAS/IS						
PL-GEN						

Background

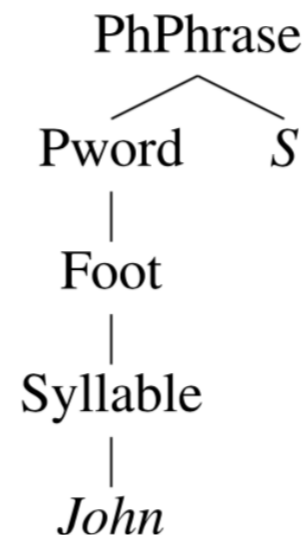
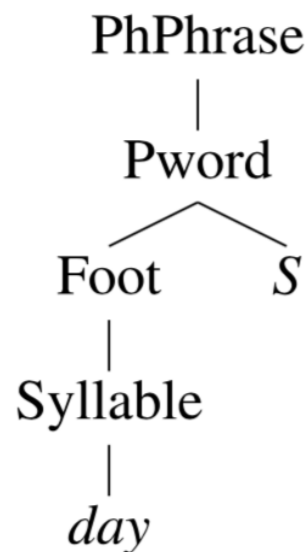
- New Zealand English (Zimmerman et al., 2016)



	S	PL	3SG	GEN	PL-G	is	has
S	///	***	***	***	***	***	***
PL		///		**	.	***	***
3SG			///		*	***	***
GEN				///		*	*
PL-G					///		
is						///	
has							///

Background: Research Questions

- Why do these differences exist?
 - Traditional explanations don't make sense (Plag et al. 2017)
 - Prosodic integration:
 - Plural S more integrated than Genitive S



- But not clear how this results in durational difference?

Background: Research Questions

- Why do these differences exist?
 - ‘discrimination management’ (Tomaschek, Plag, Ernestus & Baayen, n.d.)
 - “in speech production, prolonging part of the acoustic signal, such as S, is dysfunctional when this signal increases the discrimination problem”
 - Can be modelled using NDL
 - More on this later

Background: Research Questions

- Are these differences communicatively relevant?
 - Requires comprehension studies
 - But, register analysis may give some insight
 - Expectation: Conversational register shows more morpho-acoustic cues
- Are durational differences part of a more general acoustic reduction?
 - Look at spectral measure
- Does it even occur in Dutch?

Background

- Suffix /s/ in Dutch
 - Non-Morphemic [S]: ze heeft in een apart huiss gewoond
 - Plural [PL]: die twee kamerss
 - Possession [GEN-POSS]: en een tientje voor m'n vaderss verjaardag
 - ~~Time [GEN-TIME]: 's avondss zijn we naar de bioscoop gegaan~~
 - Partitive [PART]: daarna eten we eventueel iets makkelijks
- GEN-TIME
 - Not productive
 - More of a circumfix than a suffix

Background

- Partitive?
 - Derivational?
 - It turns an adjective into a noun (but those ‘nouns’ cannot be used elsewhere)
 - A special case of inflection used in a specific construction
- Predictions based on English
 - Non-morphemic > Plural
 - Non-morphemic > Possession (??)
 - Dutch GEN-POSS is limited to proper names and addressable nouns (e.g. *moeder*)
 - *van* ‘of’ is used more frequently
 - Alternative forms:
 - *Tim z’n fiets* ‘Tim his bicycle’
 - *Sara d’r huis* ‘Sara her house’

Data

- Natural conversations between 2 friends / acquaintances
 - CGN-A, CGN-C, CGN-D, IFADV, ECSD
 - Face-to-face or telephone conversations
- News reports
 - CGN-K
- Read-aloud stories
 - CGN-O
- Dataset that entered analysis
 - Northern Dutch (i.e. non-Flemish)
 - No overlapping speech in audio signal
 - No hesitations, incomplete words etc.
 - No /s/ followed by other sibilants
 - No /s/ with atypically long durations (> 0.4 s)

Data

- Phones were forced-aligned using CLST Forced Aligner (based on KALDI)
- Uses lexical expansion based on Schuppler et al (2011)

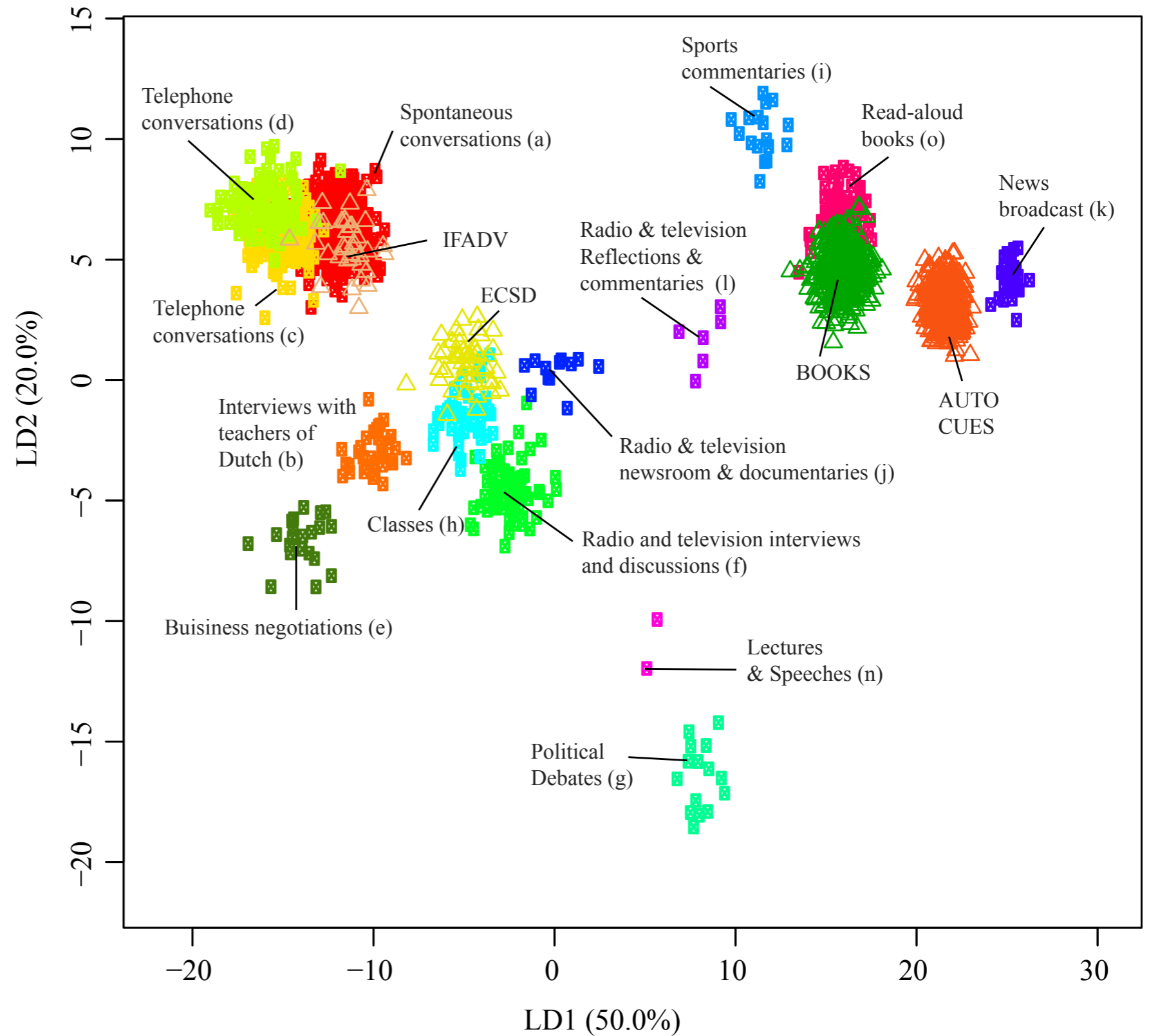
```
fietsen f i t s  
fietsen f i t s @  
fietsen f I t s @  
fietsen f i t s n  
fietsen f i t s @ n  
fietsen f I t s n  
fietsen f I t s @ n  
fietsen f @ t s @  
fietsen f @ t s n  
fietsen f @ t s @ n  
fietsen f t s n
```

Data

Register	Corpus	All /s/ Tokens	S	GEN- POSS	PART	PL
Conversation	CGN component a	84412	76934	104	873	6501
	CGN component c	36089	33531	52	311	2195
	CGN component d	24835	23018	29	225	1563
	IFADV	3751	3437	2	18	294
	ECSD (excl. negotiations)	4675	4253	2	50	370
News	CGN component k	15254	11231	43	19	3961
Stories	CGN component o	29025	23932	206	305	4582

Data

speech register classifier based on word predictability



Bentum, ten Bosch, van den Bosch, Ernestus (2019)

Analysis – Model

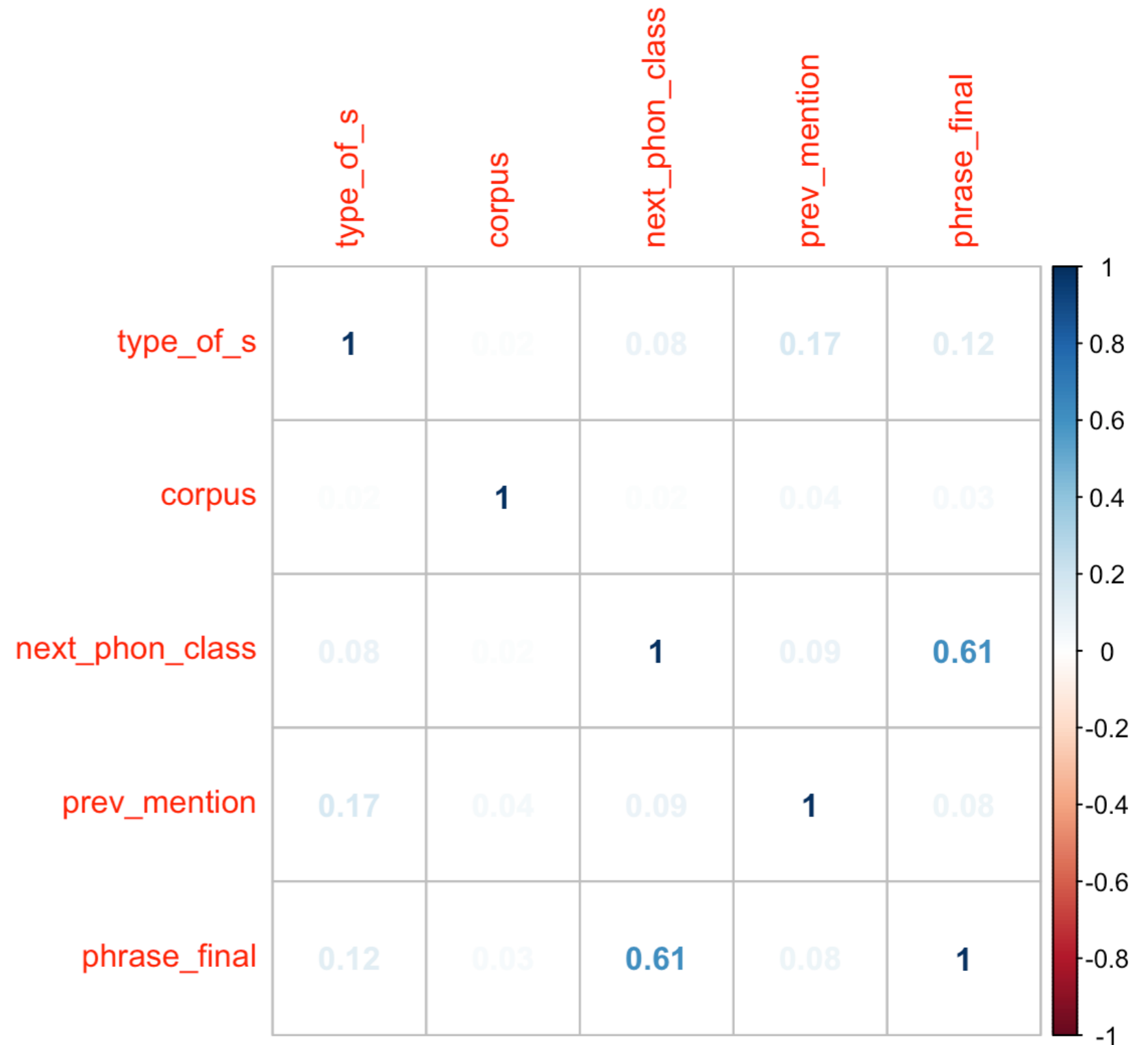
Categorical predictors

	Predictor name
Type of S	type_of_s
Following context	next_phon_class
Previous mention	prev_mention
Syntactic position	phrase_final
Corpus	corpus

Analysis – Model

Categorical predictors

- Cramèr's V
- Association between next_phon_class and phrase_final is due to silences at the end of phrases



Analysis – Model

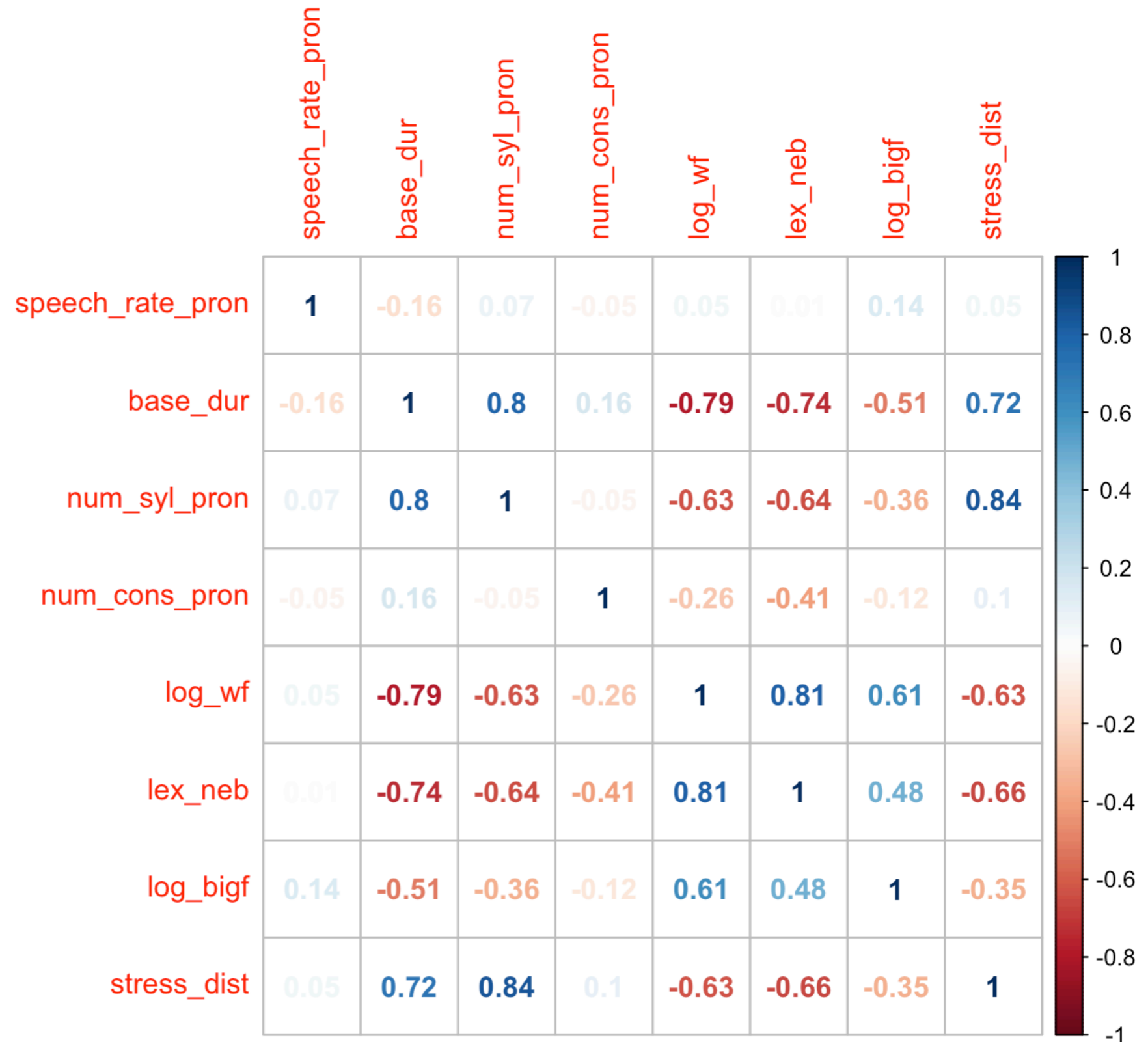
Continuous predictors

	Predictor name
Local speech rate (syl/sec)	speech_rate_pron
Base duration	mean_syl_dur
N of preceding consonants	num_cons_pron
Word frequency	log_wf
N of phon. neighbours	prop_lex_neb_freq
Bigram frequency	p_next_w
Distance to word stress	stress_dist

Analysis – Model

Continuous predictors

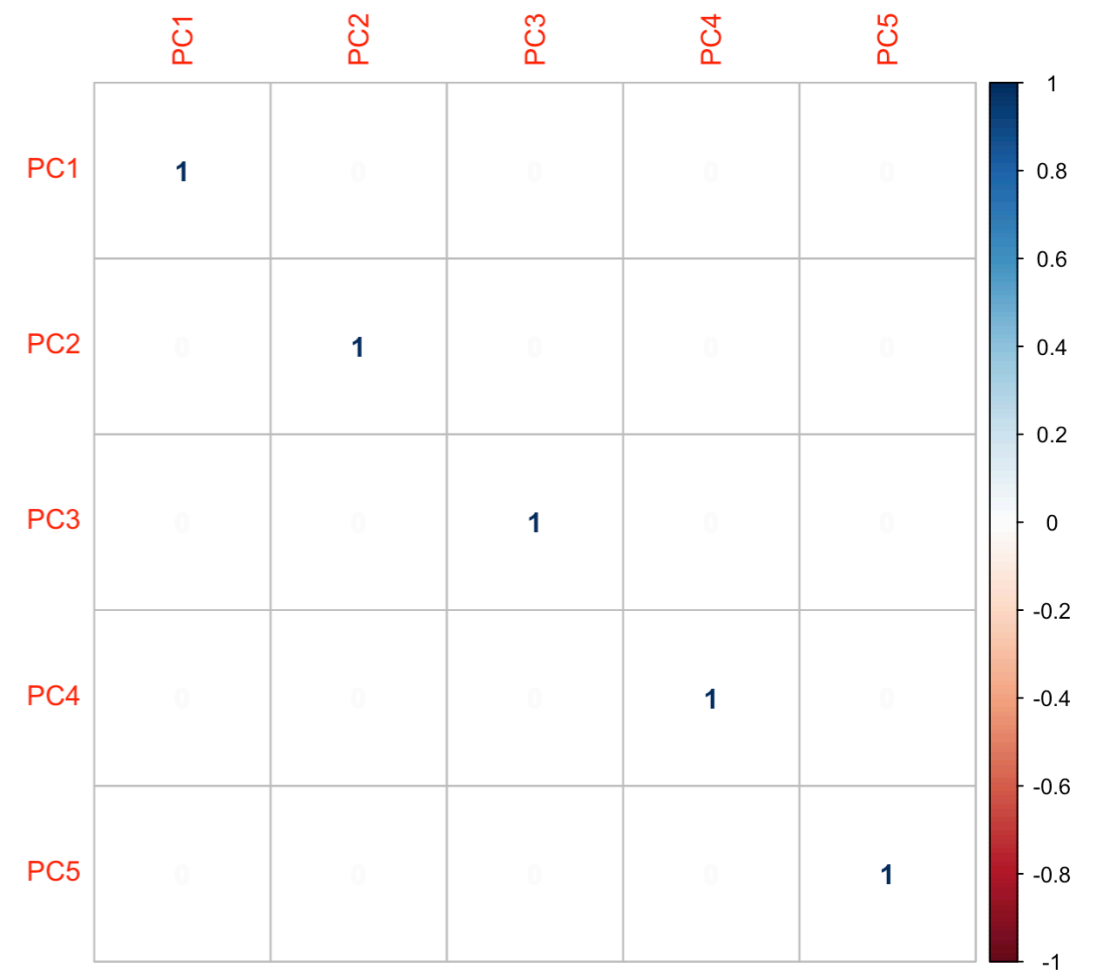
- Before PCA



Analysis – Model

PC1 contains durational & word frequency measures

predictors	PC1 Loading
base_dur	-0.44
log_wf	0.43
lex_neb	0.43
num_syl_pron	-0.41
stress_dist	-0.40
log_bigf	0.30
num_cons_pron	-0.12
speech_rate_pron	-0.0055



Analysis – Model

Categorical – Continuous associations

- Pearson's r derived from R^2 in $\text{lm}(\text{continuous} \sim \text{categorical})$



Analysis – Model

Two modelling strategies

Mixed effects model

```
s_dur ~ covariates  
+ type_of_s * register  
+ (1 | speaker) + (1 | item)
```

Mixed effects model

```
s_dur ~ covariates  
+ (1 | speaker) + (1 | item)
```

residuals

Regression model

```
resid ~ type_of_s * register
```

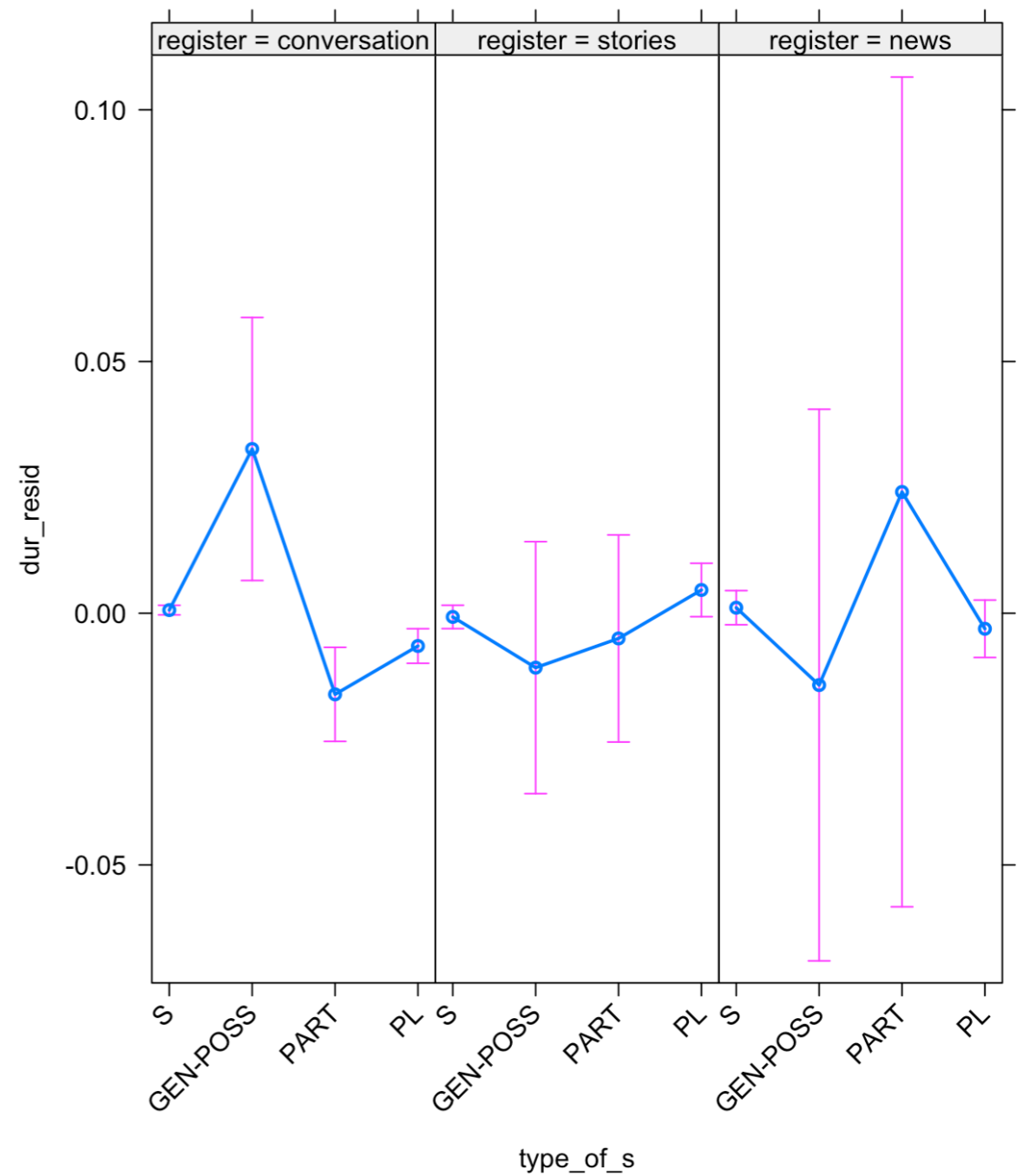
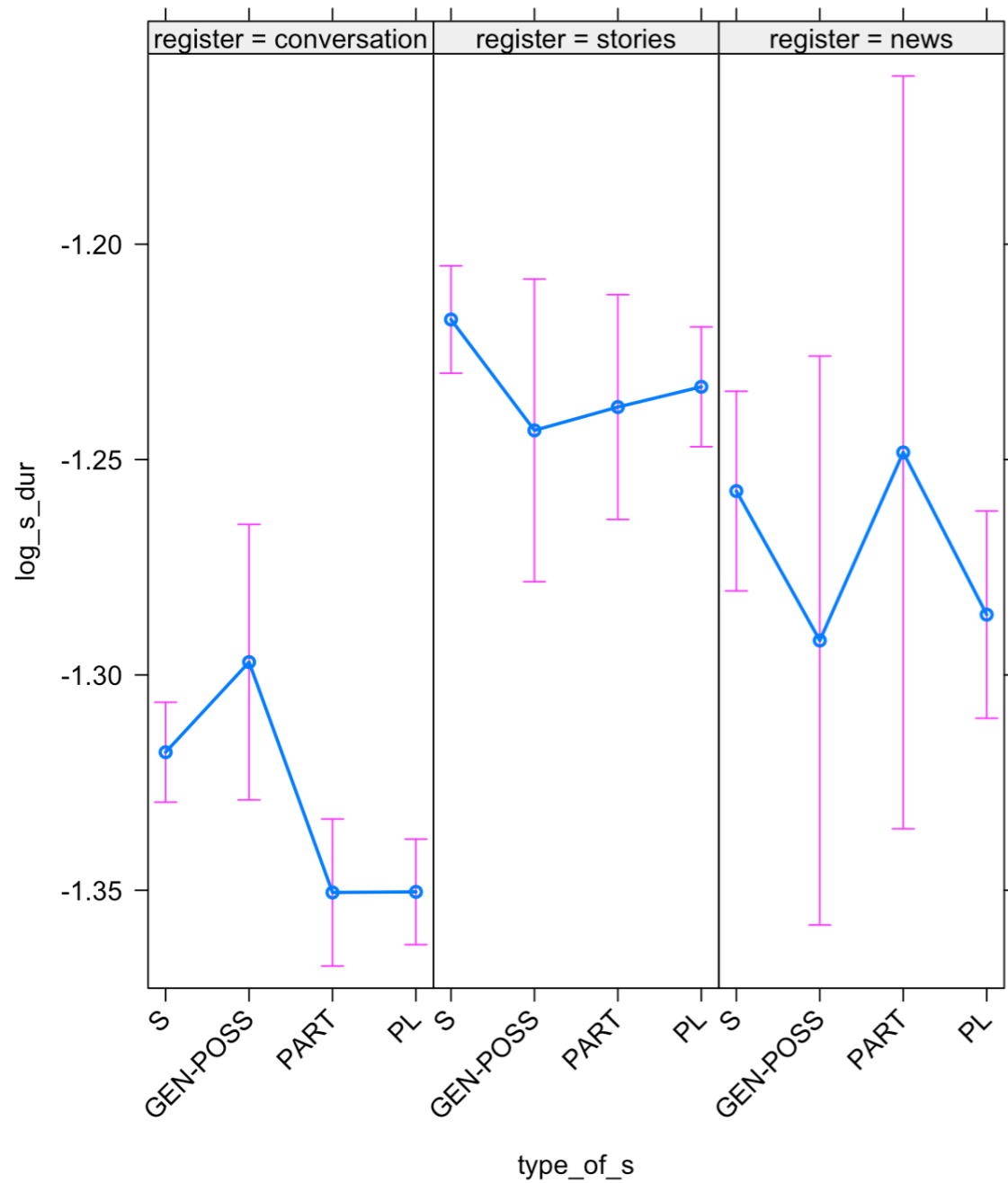
Analysis – Duration Results

Test for significance of interaction

- $s_dur \sim \dots \text{type_of_s} * \text{register}$
 - $p < 0.001$

- $\text{resid} \sim \text{type_of_s} * \text{register}$
 - $p < 0.01$

Analysis – Duration Results



Analysis – Duration Results

- Data split into separate corpora, contrasts (Tukey adjusted *p-values*)

Conversation <i>F</i> = 26.46, <i>p</i> < .001	S	GEN- POSS	PART	PL
S	X		***	***
GEN-POSS		X		**
PART			X	
PL				X
Stories <i>F</i> = 11.07, <i>p</i> < .001	S	GEN- POSS	PART	PL
S	X		***	***
GEN-POSS		X		
PART			X	
PL				X
News <i>F</i> = 0.86, <i>p</i> = .46	S	GEN- POSS	PART	PL
S	X			
GEN-POSS		X		
PART			X	
PL				X

Conversation <i>F</i> = 9.81, <i>p</i> < .001	S	GEN- POSS	PART	PL
S	X		**	**
GEN-POSS		X	**	*
PART			X	
PL				X
Stories <i>F</i> = 2.22, <i>p</i> = .08	S	GEN- POSS	PART	PL
S	X			
GEN-POSS		X		
PART			X	
PL				X
News <i>F</i> = 1.31, <i>p</i> = .27	S	GEN- POSS	PART	PL
S	X			
GEN-POSS		X		
PART			X	
PL				X

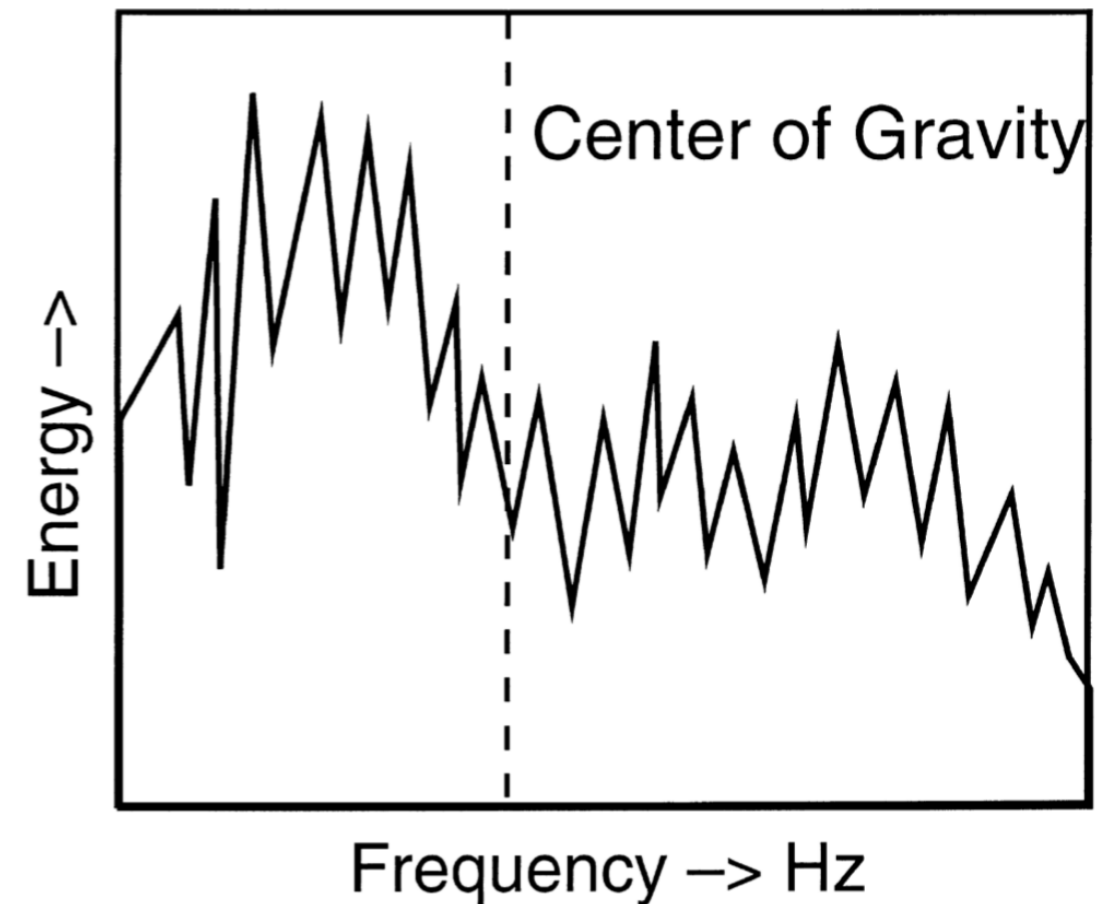
Analysis – Duration Results

Interim conclusions

- Morphology effects duration of Dutch /s/
 - Consistent with English: **S > PL**
 - New: **S > PART**
- GEN > PL/PART? (but... some FA issues for GEN)
- Non-conversational registers do not consistently show these effects

Analysis – CoG Results

- Centre of Gravity
 - “average frequency (dashed line) weighted by the acoustic power (energy)”
- If durational effects reflect reduction
 - Reduced segments should have lower CoG



van Son, Pols (1999)

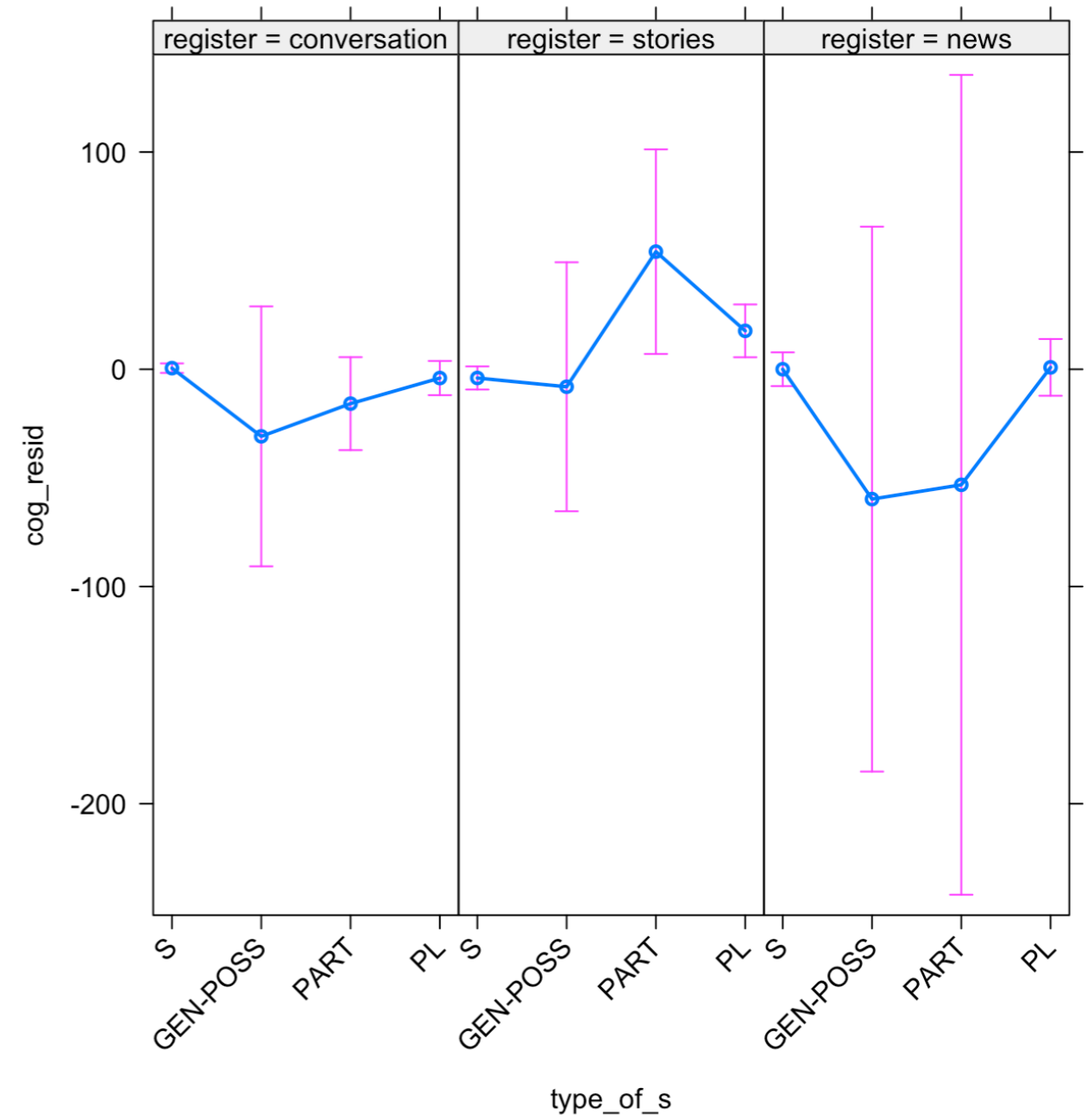
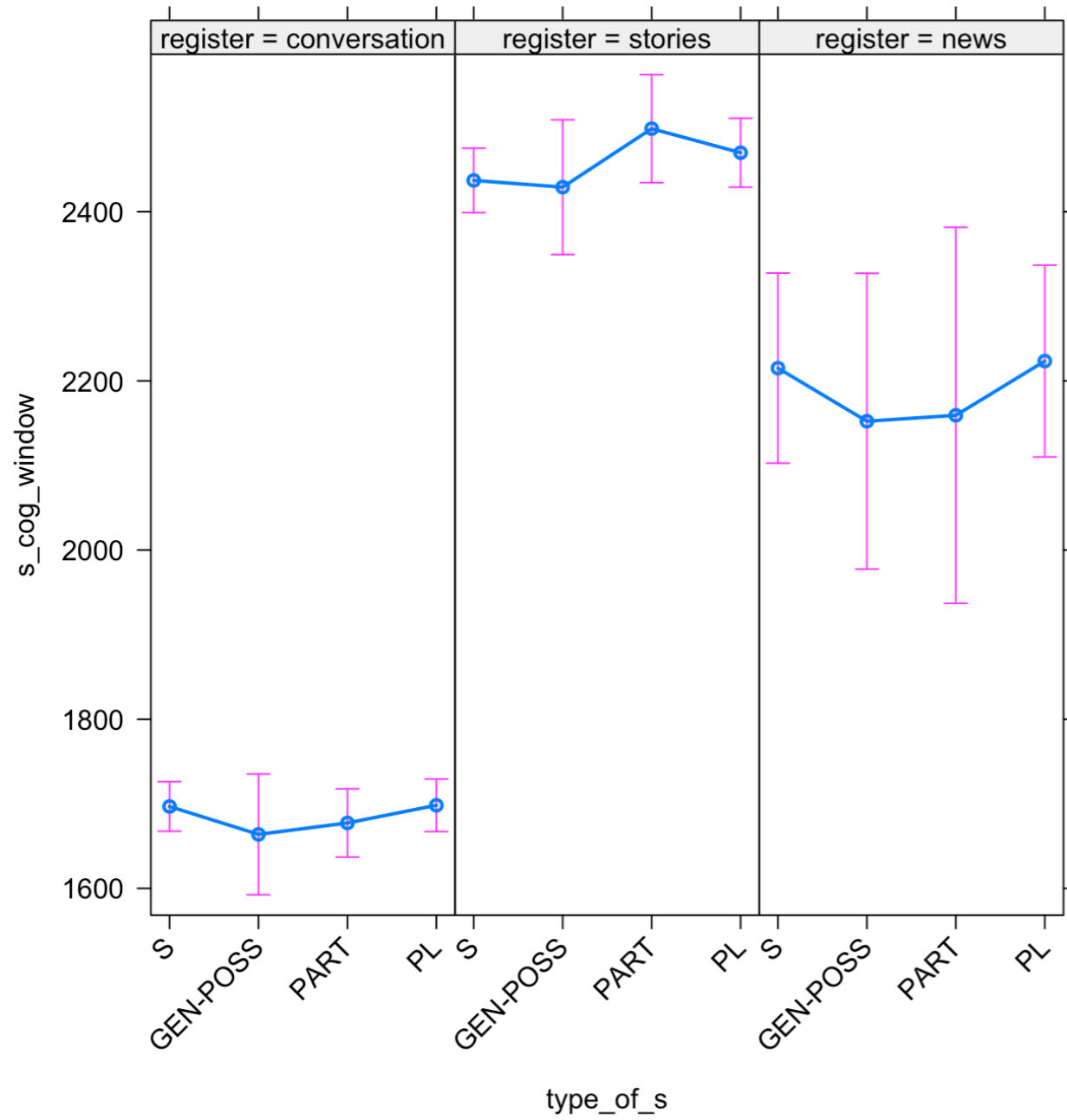
Analysis – CoG Results

Test for significance of interaction

- $s_cog \sim \dots \text{type_of_s} * \text{register}$
 - $p < 0.01$

- $\text{resid} \sim \text{type_of_s} * \text{register}$
 - $p < 0.01$

Analysis – CoG Results



Analysis – CoG Results

- Data split into separate corpora, contrasts (Tukey adjusted *p-values*)

Conversation <i>F</i> = 0.84, <i>p</i> = .47	S	GEN- POSS	PART	PL
S	X			
GEN-POSS		X		
PART			X	
PL				X
Stories <i>F</i> = 0.79, <i>p</i> < .50	S	GEN- POSS	PART	PL
S	X			
GEN-POSS		X		
PART			X	
PL				X
News <i>F</i> = 8.10, <i>p</i> < .001	S	GEN- POSS	PART	PL
S	X			***
GEN-POSS		X		
PART			X	
PL				X

Conversation <i>F</i> = 1.40, <i>p</i> = .24	S	GEN- POSS	PART	PL
S	X			
GEN-POSS		X		
PART			X	
PL				X
Stories <i>F</i> = 5.88, <i>p</i> < .001	S	GEN- POSS	PART	PL
S	X		.	*
GEN-POSS		X		
PART			X	
PL				X
News <i>F</i> = 0.46, <i>p</i> = .71	S	GEN- POSS	PART	PL
S	X			
GEN-POSS		X		
PART			X	
PL				X

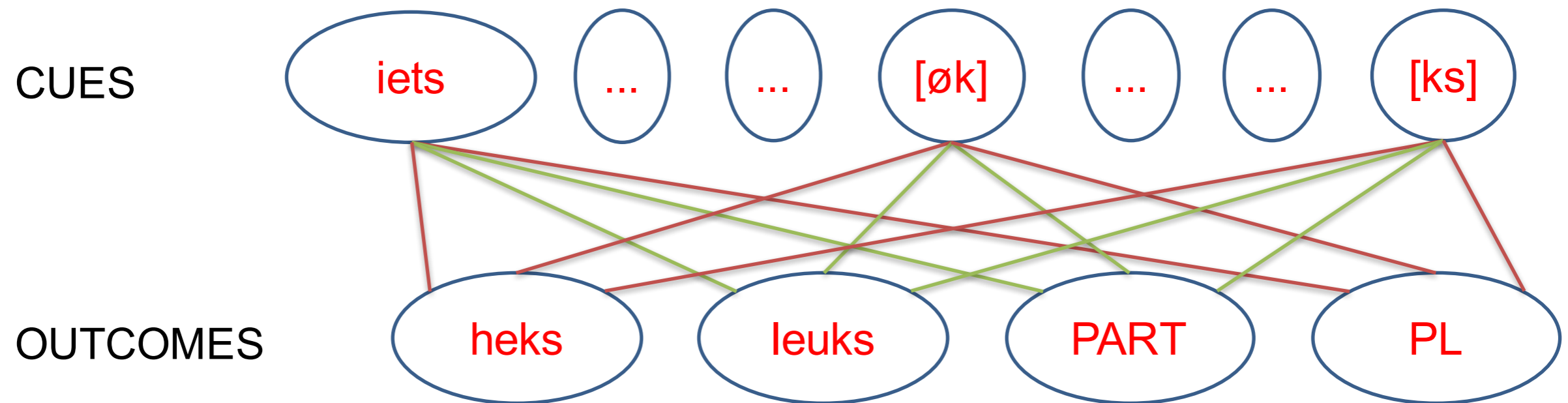
Analysis – CoG Results

Interim conclusions

- Effects inconsistent
- Effects that were found not in line with reduction

Analysis – Naive Discriminative Learning

- You encounter: “iets leuks doen” (do something fun)
 - Weights between input cues (bigrams, context words) are updated



Analysis – Naive Discriminative Learning

- If we do this for a lot of encounters we get an NDL network (a weight matrix) from which we can derive informative measures
- I trained a (small) NDL network on IFADV
- Similar cues & outcomes as Tomaschek, Plag, Ernestus & Baayen (n.d.)
 - Cues: Bigrams and ‘lexomes’ of target word, 2 preceding words and 2 subsequent words
 - Outcomes: ‘lexome’ of target word and ‘lexome’ of morphological function

Analysis – Naive Discriminative Learning

- Derived measures
- **priorMorph**: baseline activation / long term support for morph. function
- **actFromCues**: higher activation of an outcome from cues means that those cues frequently and exclusively occurred with those outcomes
- **actDivFromCues**: Higher activation diversity indicates that cues are linked to many different outcomes

Analysis – Naive Discriminative Learning

- Let's look at conversational register
- And see what remains of type_of_s effect

Mixed effects model

$s_dur \sim ndl_predictors$
 $+ (1 | speaker) + (1 | item)$

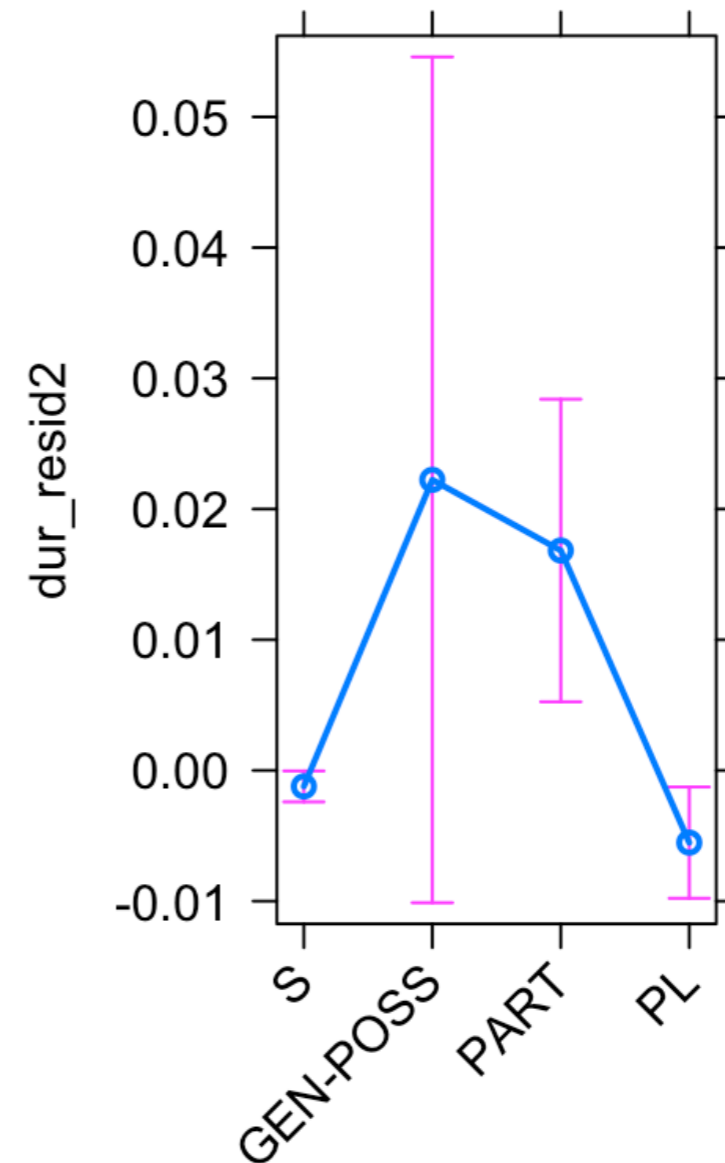
residuals

Regression model

$resid \sim type_of_s$

Analysis – Naive Discriminative Learning

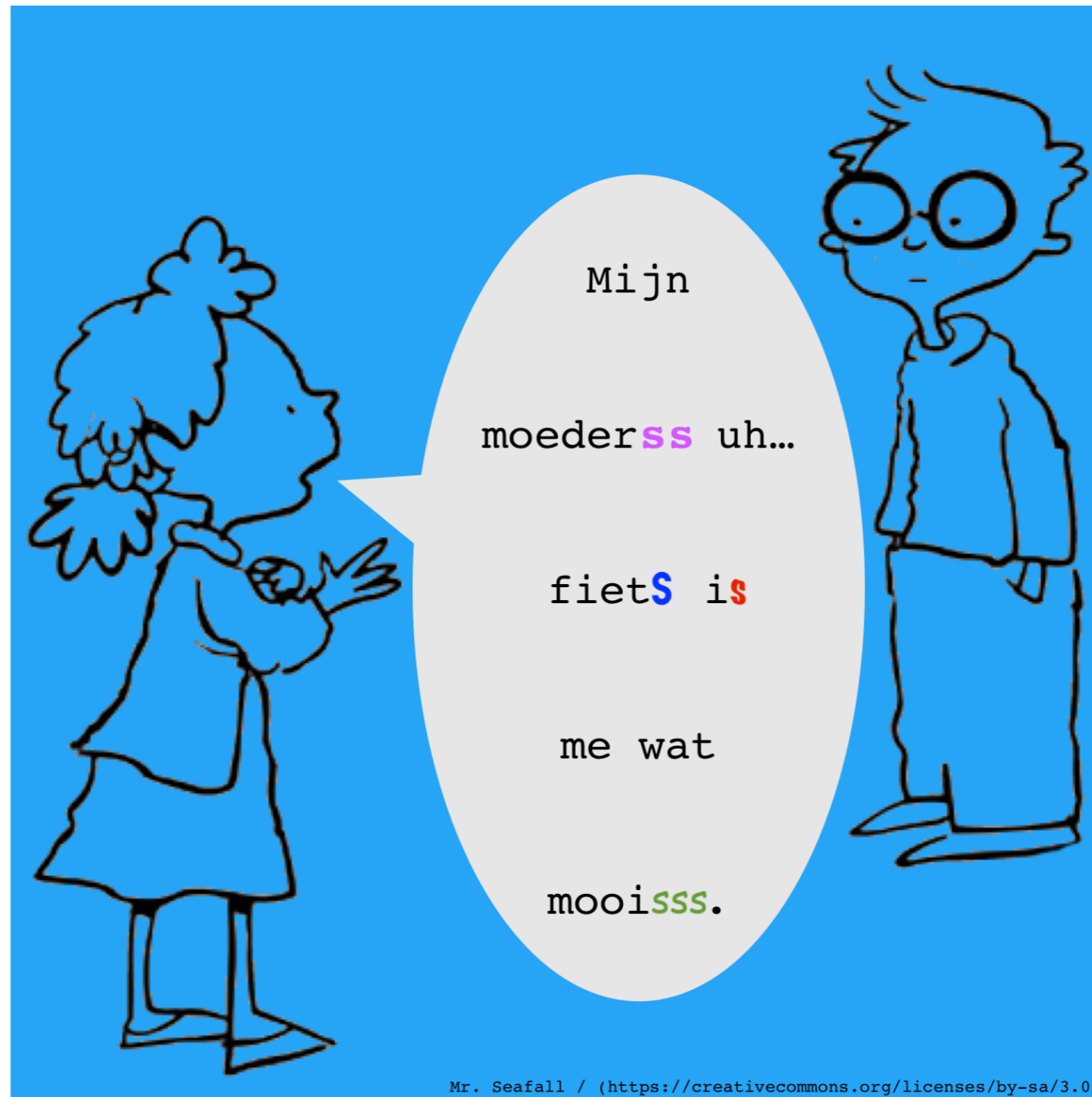
- S > PL disappears
- But NDL predictors do not have an effect that is similar to the morphological category PART
 - Not very surprising given small amount of training data for this category



Conclusions

- Production of Dutch final /s/ varies with morphological status
- Influence of register
- Preliminary results show that NDL might be a nice framework to explain differences

Questions & comments



References

- Bentum, M., Ten Bosch, L., Van den Bosch, A., & Ernestus, M. (2019). Do speech registers differ in the predictability of words?. *International Journal of Corpus Linguistics*, 24(1), 98-130.
- Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and brain sciences*, 22(1), 1-38.
- Plag, I., Homann, J., & Kunter, G. (2017). Homophony and morphology: The acoustics of word-final s in english 1. *Journal of Linguistics*, 53(1), 181-216.
- Schuppler, B., Ernestus, M., Scharenborg, O., & Boves, L. (2011). Acoustic reduction in conversational Dutch: A quantitative analysis based on automatically generated segmental transcriptions. *Journal of Phonetics*, 39(1), 96-109.
- Van Son, R. J. J. H., & Pols, L. C. (1999). An acoustic description of consonant reduction. *Speech communication*, 28(2), 125-140.
- Tomaschek, F., Plag, I., Ernestus, M., & Baayen, R. (n.d.). Phonetic effects of morphology and context: Modeling the duration of word-final S in English with naïve discriminative learning. *Journal of Linguistics*, 1-39. doi:10.1017/S0022226719000203
- Zimmermann, J., Carignan, C., & Tyler, M. D. (2016, December). Morphological status and acoustic realization: Findings from New Zealand English. In *Proceedings of the 16th Australasian International Conference on Speech Science and Technology* (pp. 6-9).