Heinrich Heine Universität Düsseldorf

Morpho-phonetics

The role of the acoustic signal in morphology

Ingo Plag

presenting joint work with

Harald Baayen, Dinah Baer-Henney, Melanie Bell, Louis ten Bosch, Marie Engemann, Mirjam Ernestus, Dominic Schmitz, Simon Stein, Fabian Tomaschek, Tim Zee

International Society of Experimental Linguistics, December 17, 2021

Morpho-phonology as we know it

1. Morpho-phonological alternations are categorical but may have lexical exceptions.

2. The formal level of representation of morphemes is phonological in nature.

3. Post-lexical phonology and phonetics have no access to lexical information.

Morpho-phonology as we know it

1. Morpho-phonological alternations are categorical but may have lexical exceptions.

2. The formal level of representation of morphemes is phonological in nature.

3. Post-lexical phonology and phonetics have no access to lexical information.

2. The formal level of representation of morphemes is phonological in nature.

English suffixes

-ic

/ɪk/ ~ /ɪs/

/ɪk/ → /ɪs/ | __ {*-ize, -ify, -ism, ...*}
/ɪk/ → /ɪk/ | elsewhere

/ɪk/ → [ɪs] | __ {*-ize, -ify, -ism, ...*}
/ɪk/ → [ɪk] | elsewhere

No reference to sub-phonemic detail

plural, 3sg, genitive

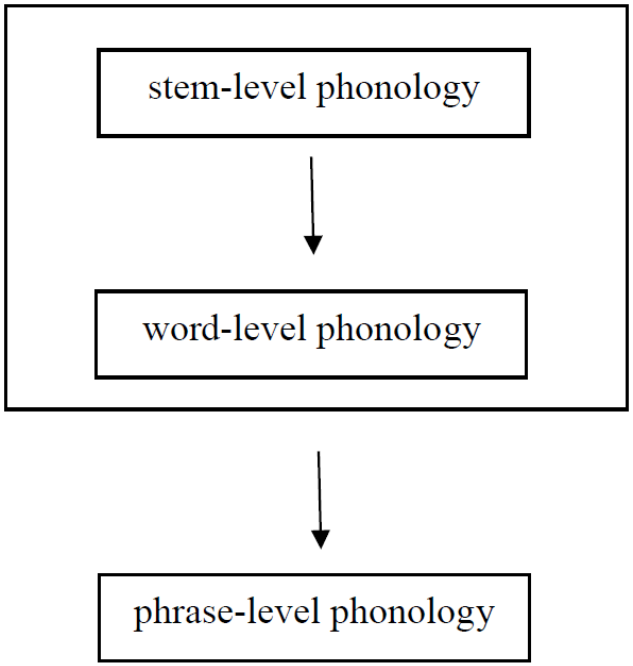
/z/ ~ /s/ ~ /ɪz/

/z/ → /s/ | ...
/z/ → /z/ | ...
/z/ → /ɪz/ |...

/z/ → [s] | ...
/z/ → [z] | ...
/z/ → [ɪz] |...

No reference to sub-phonemic detail

3. Post-lexical phonology and phonetics have no access to lexical information.



LEXICON

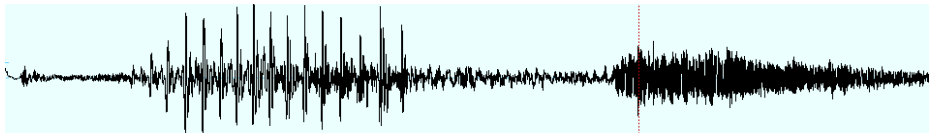
SYNTAX

(from Plag (2018) *Word-formation in English 2nd ed.* CUP, chapter 7)

Speech production

(Levelt et al. 1999, Roelofs & Ferreira 2019)

Concept

Lemma	‘cap‘	Noun [+ concrete] [+count]	‘more than one‘ [plural]
Phonological representation	/kæp/		/-z/
(Morpho-)phonological rules		[kʰæps]	
Syllabary		[$ kæps $]	
Articulation		<movement of articulators>	
Acoustic signal			

Crucial point

No morphological information available post-lexically

Problem: Morpho-phonetics

Reasons to worry

- Free and bound variants of a base differ acoustically (Kemps et al. 2005, Blazej & Cohen-Goldberg 2015)
- Duration of Dutch compound linking morphemes depends on paradigmatic probability (Kuperman et al. 2007)
- Vowel frontness of Russian verbal suffix depends on paradigmatic probability (Cohen 2014a)
- Duration of English 3sg S depends on syntagmatic probability (Cohen 2014b)
- Duration of English final S depends on its morphological status (Zimmermann, 2016, Plag et al. 2017, Plag et al. 2020, Schmitz et al. 2021)
- ...

Challenge for existing theories of morpho-phonology, the mental lexicon and speech production

1. What morpho-phonetic effects are observable? Which morphological structures show morpho-phonetic effects?
2. How do these effects play out?
3. How can these effects be explained?

Today

- Some recent findings

Which morphological structures show which morpho-phonetic (acoustic) effects?

Stems

- Duration: free vs. bound stems, stems with different affixes (Kemps et al. 2005, Blazey & Cohen 2015, Cohen 2014b, Plag et al. 2020, Zee et al. 2021, Stein & Plag 2021, etc.)
- Aspiration of stem-initial plosives after prefixes (Smith et al. 2012, Zuraw et al. 2020, Popescu 2021)

Affixes

- Duration: suffixes, prefixes (Zimmermann 2016, Plag et al. 2017, Seyfarth et al. 2017, Plag & Ben Hedia 2019, Schmitz et al. 2021)

Paradigms

- Duration: linking morphemes, stems, suffixes (Kuperman et al. , Seyfarth et al. 2017, Plag et al. 2020, Engemann & Plag 2020, Zee et al. 2021)
- Formants: Vowel frontness (Cohen 2014a)

Boundaries

- Duration: compounds, prefixes, suffixes (Ben Hedia & Plag 2018, Ben Hedia 2019, Bell et al. 2020)
- Coarticulation/assimilation across morphological boundaries (Cho 2008, Tomaschek et al. 2020, Saito 2020)

Null results

- English final *-s* vs. *-ed* (Zimmermann 2016, Seyfarth et al. 2017)
- English plural *-s* vs. pluralia tantum *-s* (Schlechtweg & Corbett 2021)
- Coarticulation across morphemic boundaries (Seyfarth et al. 2018, Mousikou et al. 2021)

How do these effects play out?
How can these effects be explained?

- Morpho-phonetic effects in paradigms
  - Engemann & Plag 2020
  - Bell, Ben Hedia & Plag 2020
  - Zee, ten Bosch, Plag & Ernestus 2021

- Discriminative Learning
  - Tomaschek, Plag, Baayen & Ernestus 2019
  - Schmitz, Plag, Baer-Henney & Stein 2021
  - Stein & Plag 2021

- Methodological lesson
  - real speech (corpora)
  - experiments
  - computational modeling

U. Marie Engemann & Ingo Plag

Plurals vs. mono-morphemic words

(*The Mental Lexicon* 2021*)*

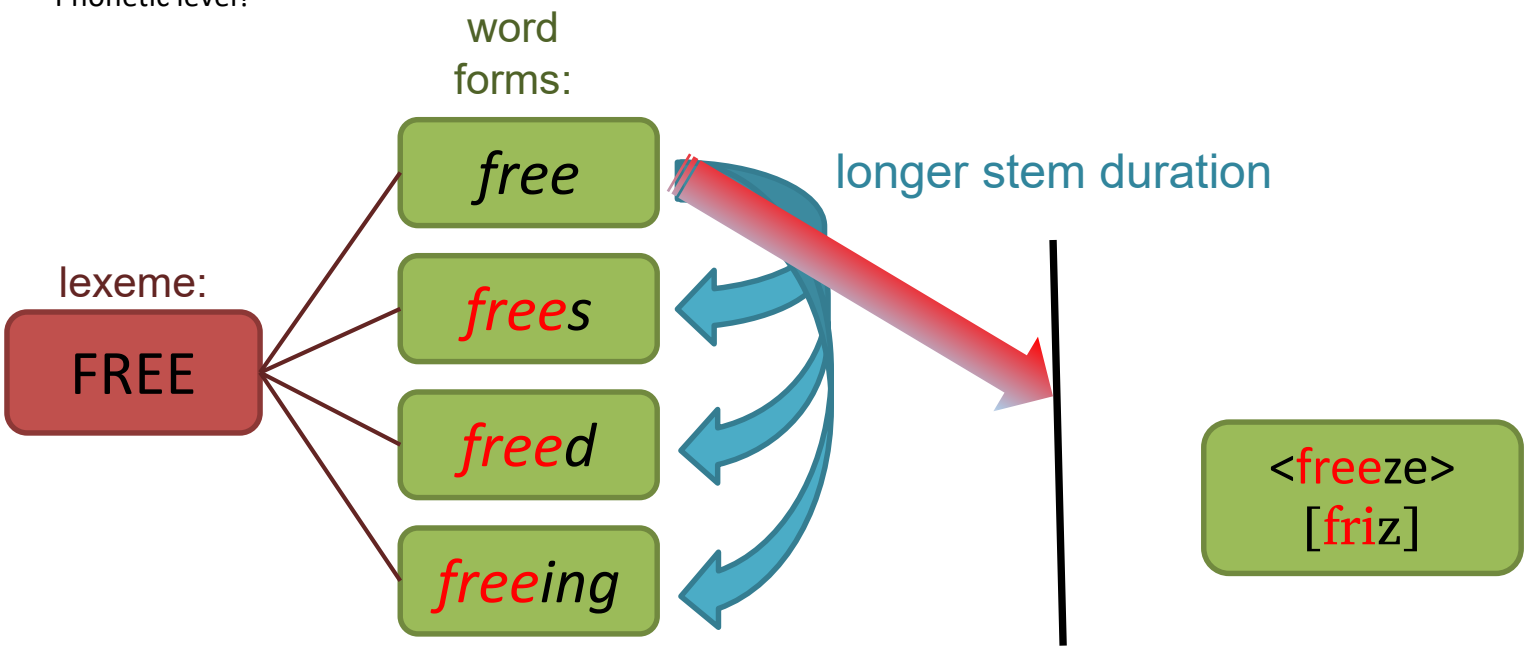
Paradigm Uniformity / paradigm leveling

- Forms in a paradigm influence each other, may become similar to each other
- Phonological level:

  stride - strode - stridden ~ stride - strode - strode (Bauer, Lieber & Plag 2013: 77)

  sólid, húmid ~ solídify, humídify ~ solídity, humídity (Plag 1999:203, Bauer, Lieber & Plag 2013: 271f)

- Phonetic level?



- Seyfarth et al. (2017) *Lang., Cogn. and Neurosc.*
  Stems of S-suffixed words are longer than the corresponding string in homophonous monomorphemic words

Our study

- Replicate the phonetic Paradigm Uniformity effect with authentic speech
- Investigate the effect of frequency / reduction
- QuakeBox Corpus (Walsh et al. 2013), NZ English, data set from Zimmermann (2016)
- Sample: Words that are
  - monosyllabic and end in /z/
  - monomorphemic or plural
  - have final /z/ preceded by a vowel
- At least 3 tokens per type: 431 tokens, 38 types

Hypotheses

H1 Stems of plural words are longer than corresponding parts of monomorphemic words (***day****s* > ***da****ze*)

Categorical Paradigm Uniformity

H2 The more frequent the bare stem (e.g. *day*), the shorter the bare stem (general reduction), hence the shorter the plural stem (as in ***day****s*)

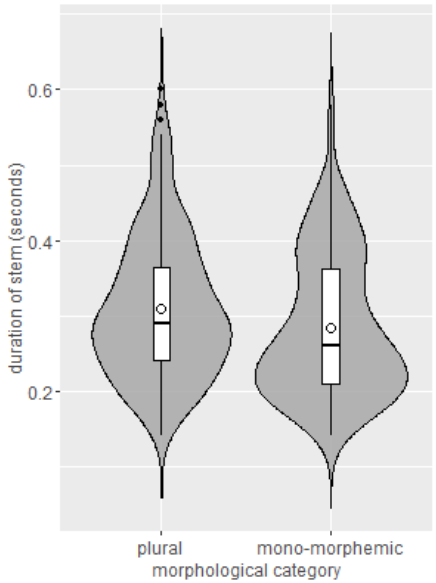
Gradient Paradigm Uniformity

Statistical modeling

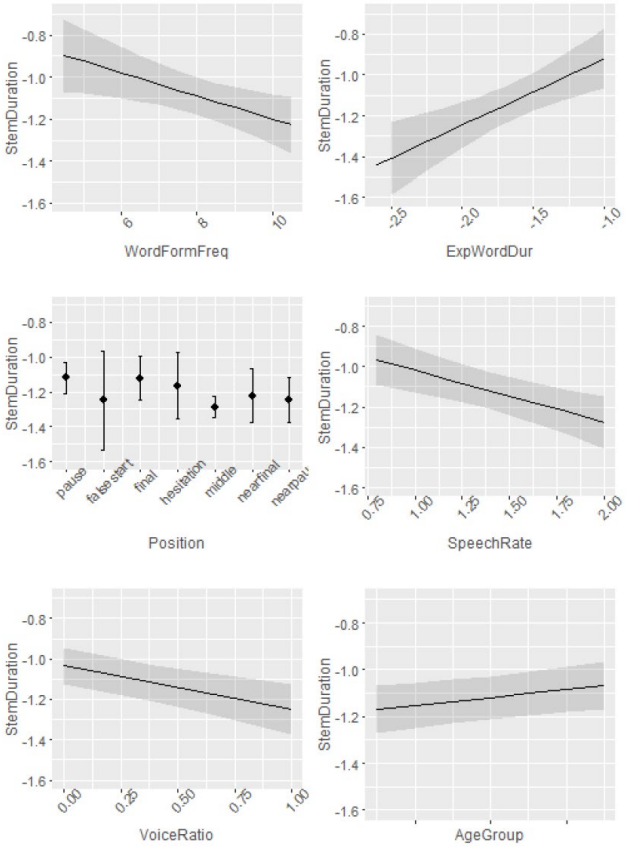
- Multiple linear mixed effects regression
- Random forests (collinearity issues)
- Variables of interest
  - MORPHEME TYPE: suffixed vs. mono-morphemic
  - WORD FORM FREQUENCY
  - STEM FREQUENCY
  - RELATIVE FREQUENCY
- Control variables
  - NUMBER of PHONEMES
  - EXPECTED STEM DURATION
  - SPEECH RATE
  - BIGRAM PROBABILITY (left and right)
  - NEIGHBORHOOD DENSITY
  - NEIGHBORHOOD FREQUENCY
  - POSITION in utterance
  - AGE GROUP
  - VOICE RATIO

Results: Categorical paradigm uniformity

H1 Stems of plural words are longer than corresponding parts of monomorphemic words (Categorical Paradigm Uniformity Effect)



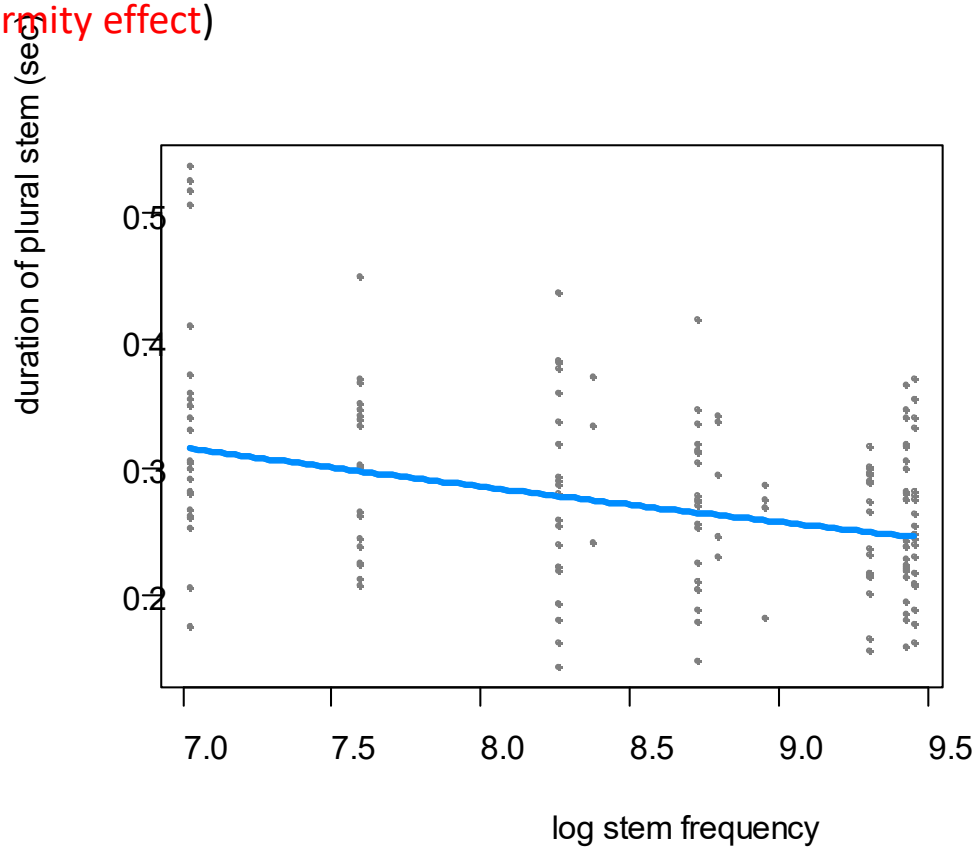
30 ms difference
not significant (*p*=0.25, t=*t*=-1.172, *df*=28.59, *strd. error*=0.059)



No Categorical Paradigm Uniformity Effect (H1 is rejected)

Results: Reduction and gradient paradigm uniformity

H2 The more frequent the bare stem, the shorter the bare stem (general reduction effect), and, due to paradigm uniformity, the shorter the plural stem (Gradient Paradigm Uniformity effect)



Significant effect of stem frequency in the predicted direction (t=-3.07, p=0.03)

The more frequent the bare stems, the shorter the plural stems

Summary: Paradigm Uniformity

- No robust effect of Categorical Paradigm Uniformity in authentic speech
- Clear effect of Gradient Paradigm Uniformity, in spite of small data set
- But how does this go together?
- Which forms in a paradigm may influence which other forms, and how do they do this?

Melanie Bell, Sonia Ben Hedia & Ingo Plag

Phonetic realization of consonants at the internal boundary in English compound nouns

(*Morphology* 2020*)*

Research question

- What factors influence the duration of consonant length at compound boundaries (e.g. *he**n** night, stea**m** engine, tuna **s**andwich*)?

- Three hypotheses:
  - Higher Informativity leads to lengthening
    (Jurafsky et al. 2001; van Son & Pols 2003)
  - Higher segmentability leads to lengthening
    (Hay 2004; Ben Hedia & Plag 2017; Plag & Ben Hedia 2018)
  - More paradigmatic support leads to lengthening
    (Kuperman et al. 2007; Cohen 2014)

- How measure informativity, segmentability and paradigmatic support in compounds?
  - Transitional probability of segments at the boundary: *he**n_n**ight, stea**m_e**ngine, tun**a_s**andwich*
  - Constituent frequency
  - Constituent families

Compound constituent families

N1 family

time machine

time limit

time period

time scale

time series

time limits

time management

time interval

time zone

N2 family

time machine

washing machine

sewing machine

knitting machine

fax machine

coffee machine

drum machine

war machine

Conflicting predictions

Conflict 1

Segmentability: greater N1 family size
-> more productive N1
-> more decomposable
-> longer consonant duration at boundary

Paradigmatic support: greater N1 family size
–> more possible values for N2
–> more spread activation/ greater uncertainty
–> shorter consonant duration at boundary

Conflict 2

Informativity: greater paradigmatic probability of consonant
–> less informative
–> shorter consonant duration at boundary

Paradigmatic support: greater paradigmatic probability of consonant
–> more support
–> longer consonant duration at boundary

Experiment

consonant	CC	CV	VC
/m/	*cream mini*	*steam engine*	*survey manager*
/n/	*hen night*	*line item*	*nursery nurse*
/s/	*class survey*	*police act*	*tuna sandwich*

- Items in carrier sentences (two types)

  They talked about the [line item] again.

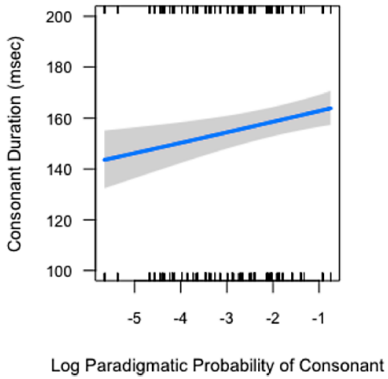
  She told me about the [tuna sandwich].

- Mixed effects regression with **consonant duration** as dependent variable

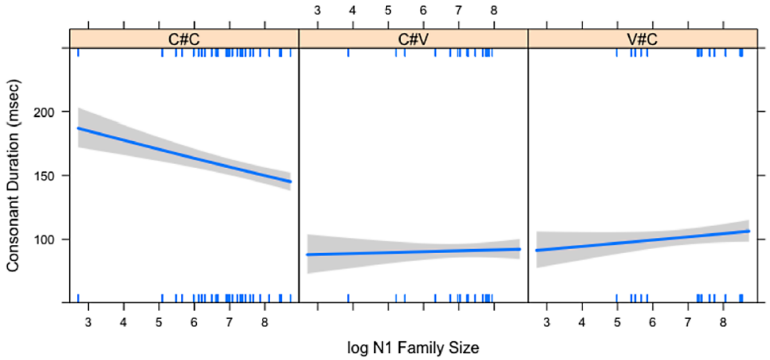
Variables of Interest

- All measures based on ukWaC (> 2bn words from the .uk internet domain)
  - Constituent family sizes
  - Paradigmatic probability of consonant at the boundary
  - Constituent frequencies
  - Spelling ratio
  - Compound frequency
  - Conditional frequency N2 given N1
  - Entropy of constituent families
- Many noise variables

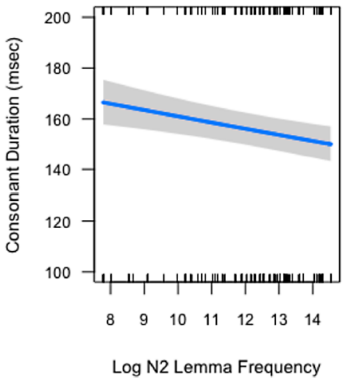
Results



- Evidence for Paradigmatic support
- Against informativity



- Evidence for Paradigmatic support
- No evidence for informativity or segmentability



- Only piece of evidence for informativity effect

Conclusion

- Paradigmatic Enhancement Hypothesis is well supported
- No support for the segmentability hypothesis
- Very restricted role of informativity

Tim Zee, Mirjam Ernestus, Louis ten Bosch & Ingo Plag

Dutch variable plurals

(Frontiers in Psychology 2021*)*

Dutch plurals: *-s* vs. *-en*

Speaker:
- Huub Pragt
- Egyptologist

Wilde Haren de podcast (2019)

en	uh	hij	beweerde	dat	die-	dat	de	**piramides**	van	Gizeh,	hè,	de	drie	bekende	**piramiden,**	dat	die ...
and	uhm	he	claimed	that	those-	that	the	**pyramids**	of	Giza,	right,	the	three	famous	**pyramids,**	that	they ...

- Choice is governed by probabilistic phonological rules (= analogy / similarity with other words)
- For some words the rules predict two possible forms

Research questions

- Does the duration of plural *-s* depend on the proportion of plural vs. singular?

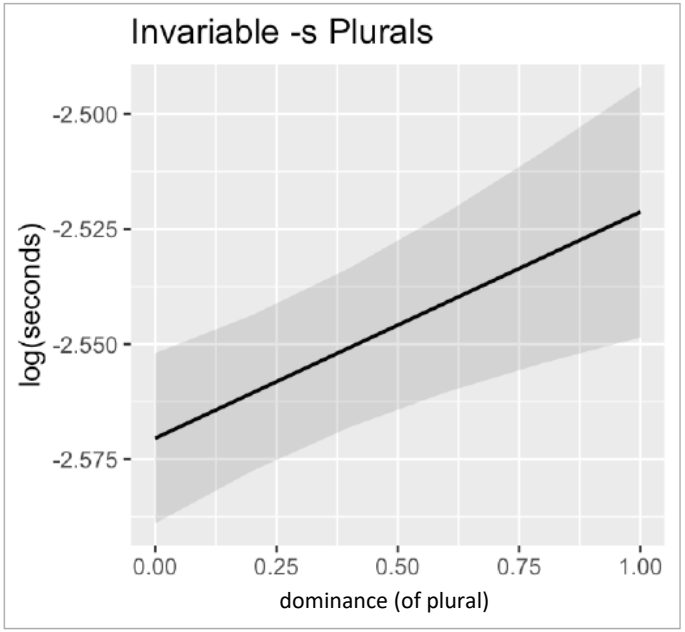
singular-dominant:	*boek - boeken*	`book(s)'
plural-dominant:	*schoen - schoenen*	`shoe(s)'

- Does the duration of plural *-s* in variable plurals (*piramides* vs. *piramiden*) depend on the proportion of *-s* plurals?

- Data: Corpora of spoken Dutch

Results for dominance

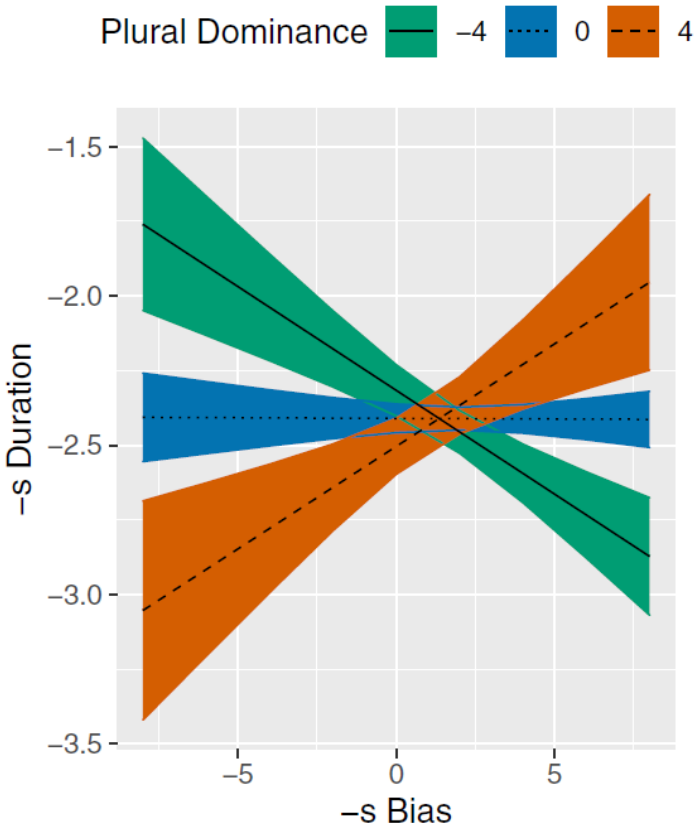
- Does the duration of plural *-s* depend on the proportion of plural vs. singular word-forms (`dominance‘)?



- Paradigmatic enhancement effect
  (cf. Dutch linking morphemes, Kuperman et al. 2007, English compounds, Bell et al. 2020)

Results for variable plurals

- Does the duration of plural *-s* in variable plurals (*piramides* vs. *piramiden*) depend on both the proportion of plurals and the proportion of *-s* plurals?



- Plural-dominant words: Paradigmatic enhancement
- Singular-dominant words: Segmentability is highest with lowest -s bias
- Singular-dominant words: Syntagmatic reduction?

Interim summary

- Co-activation of morphologically related forms in paradigms influences speech production.
- Paradigmatic enhancement
  Higher predictability of a form within its paradigm leads to longer duration.
- Reduction
  Paradigm-related reduction effects are also observed.
- The presence of such effects is unexplained in standard models speech production.
- The relation between reduction effects and enhancement effects is not totally clear.

Towards explanations: Discriminative learning

Naive 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)

- Learning results from exposure to informative relations among events in the environment (co-occurrence of cues and outcomes), partly 'error-driven'

- Association weights, adjusted according to new, informative experiences ('Rescorla-Wagner equations')

- Association weights ≈ contextual and paradigmatic predictability

- General idea
  Association weights may be used to predict language behavior

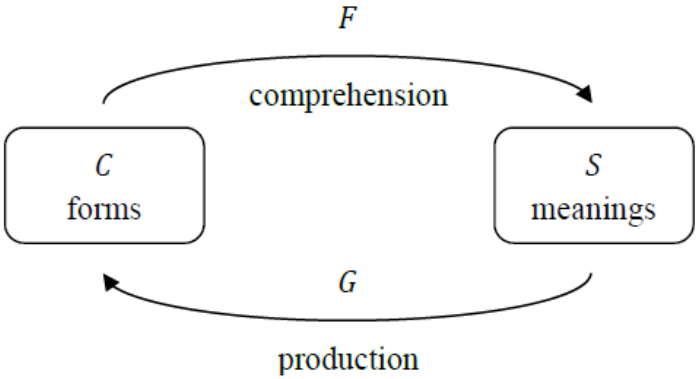
Linear Discriminative Learning

(Baayen and colleagues 2019 et seq.)

- LDL networks: simple two-layer networks
- Linguistically transparent and interpretable
- Mental lexicon: five high-dimensional numeric matrices
- Five different subsystems:

visual matrix	⇔	retina
auditory matrix	⇔	cochlea
semantic matrix	⇔	mental representations of meaning
form matrix	⇔	speaking
spelling matrix	⇔	typing

- Linear mapping of matrices: form onto meaning, meaning onto form



Mapping form and meaning

$$T = \begin{array}{c} \\ cat \\ bus \\ eel \end{array} \begin{array}{c} \begin{array}{cccccccc} \#k\{ & k\{t & \{t\# & \#bV & bVs & Vs\# & \#il & il\# \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \end{array}$$

$$S = \begin{array}{c} \\ cat \\ bus \\ eel \end{array} \begin{array}{c} \begin{array}{ccc} cat & bus & eel \end{array} \\ \begin{pmatrix} 1.0 & 0.2 & 0.5 \\ 0.4 & 1.0 & 0.1 \\ 0.2 & 0.3 & 1.0 \end{pmatrix} \end{array} .$$

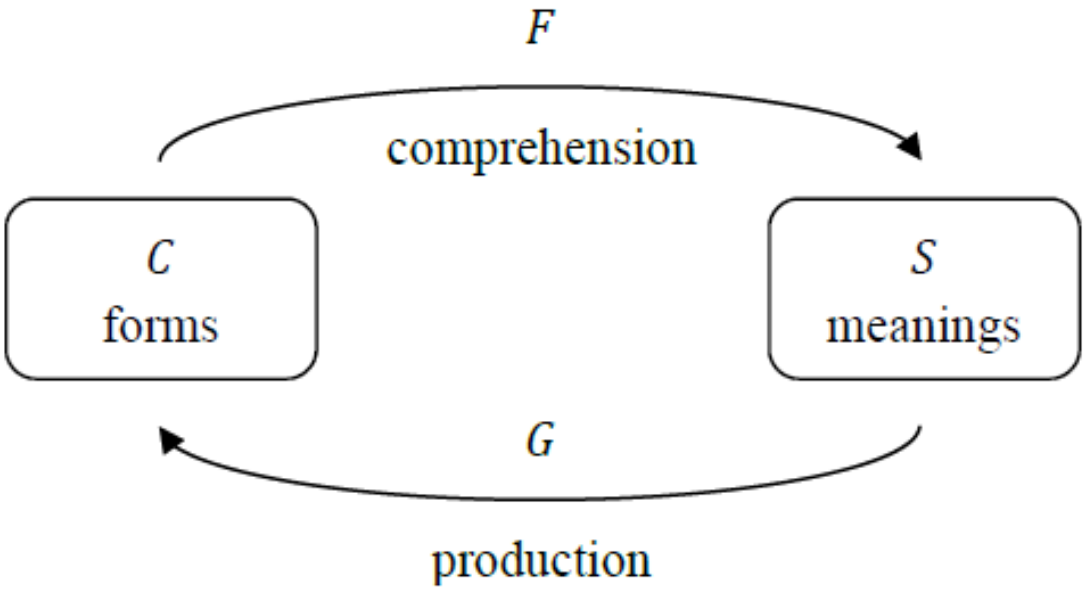
- Transformation matrices *F* and *G* can be used to predict meanings and forms

$$S = F * T \qquad T = G * S$$

$$F = S' * T \qquad G = T' * S$$

$$\hat{S} = F * T \qquad \hat{T} = G * S$$

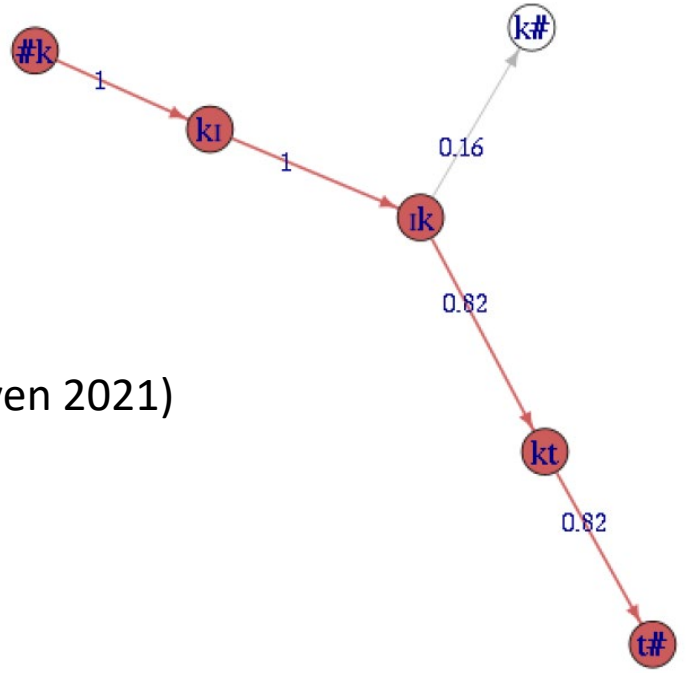
Predicting form, predicting meaning



Prediction of a particular form

- Predict the form matrix on the basis of the semantic matrix and the transformation matrix *G*
- Predicted form matrix: amount of support for individual triphones by the semantics of the word (e.g. Baayen et al. 2018)

	#vo	voc	oco	coo	oo#	oca	caa	aas
vocoo	1.00	1.00	0.44	0.42	0.50	0.44	0.49	-0.02
vocaas	1.01	1.01	0.06	0.04	0.04	0.91	0.90	0.43
vocat	1.00	1.00	0.00	0.02	0.02	0.97	0.50	-0.02



- linking of highly supported triphones
  (using graph theory, e.g. Chuang & Baayen 2021)

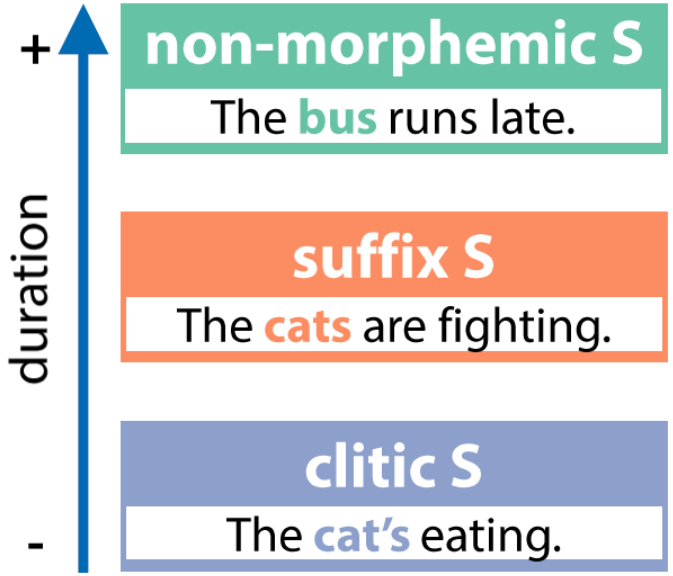
Discriminative learning and phonetics

- Idea: Use measures from the discriminative networks to predict duration
- NDL implementation
  Duration of final *-s* in English speech corpus (Buckeye)
  Tomaschek at al. 2019, *Journal of Linguistics*
- LDL implementation
  Duration of final *-s* in English pseudowords (experiment)
  Schmitz et al. 2021, *Frontiers in Psychology*
- LDL implementation
  Duration of derived words in English (BNCAudio)
  Stein & Plag 2021, *Frontiers in Psychology*

Final -*s*

Tomaschek et al. 2019, Schmitz et al. 2021

- Zimmermann (2016)
- Plag et al. (2017)
- Tomaschek et al. (2019)
- Schmitz et al. (2020) on nonce words



Results

Tomaschek et al. 2019 (data from Buckeye Corpus):

- Increasing support for a morphological function from the network goes together with longer durations.
- Higher activation diversity for the different morphological functions (≈lower paradigmatic predictability) leads to shorter /s/ durations

Schmitz et al. 2021 (data from pseudo-word experiment)

- Higher semantic support for a particular phonological path leads to shorter duration
- Higher semantic activation diversity lead to shorter S durations

Duration of derived words and their consitutents

Stein & Plag 2021

- What determines the duration of derived words and their constituents?
- Inconclusive results for segmentability, informativity, prosody

	tokens	*types*	*derivational functions*
BNC Audio	4530	363	DIS, NESS, LESS, ATION, IZE

Duration in derived words: LDL predictors

predictor	represents	defined as
MEAN WORD SUPPORT	articulatory certainty	$\frac{sum\ of\ path\ supports}{number\ of\ path\ nodes}$

Duration in derived words: LDL predictors

predictor	represents	defined as
MEAN WORD SUPPORT	articulatory certainty	$\frac{sum\ of\ path\ supports}{number\ of\ path\ nodes}$
PATH ENTROPIES	articulatory uncertainty	Shannon entropy of path supports

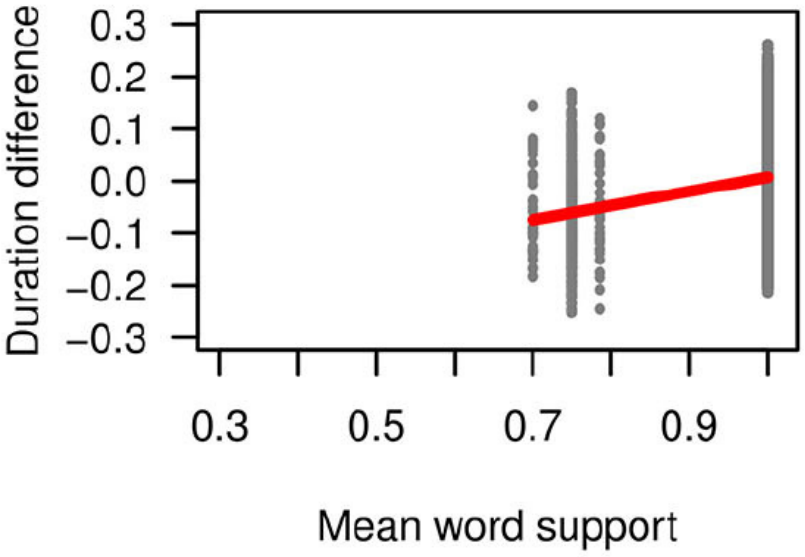
Duration in derived words: LDL predictors

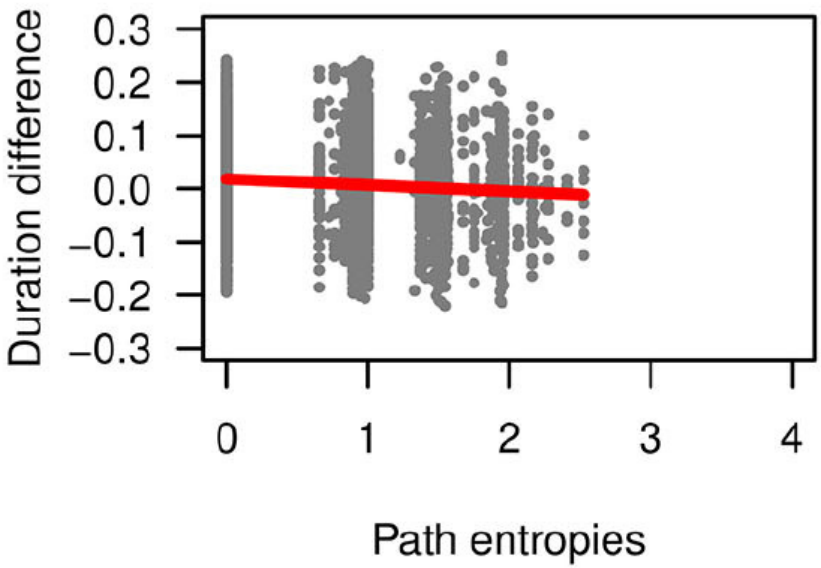
predictor	represents	defined as
MEAN WORD SUPPORT	articulatory certainty	$\frac{sum\ of\ path\ supports}{number\ of\ path\ nodes}$
PATH ENTROPIES	articulatory uncertainty	Shannon entropy of path supports
SEMANTIC VECTOR LENGTH	semantic activation diversity	mean correlation of $\hat{s}$ with top 8 neighbors

Dependent variable: 'Duration difference'

Difference in duration between the expected duration (sum of average sgement durations) and the observed duration

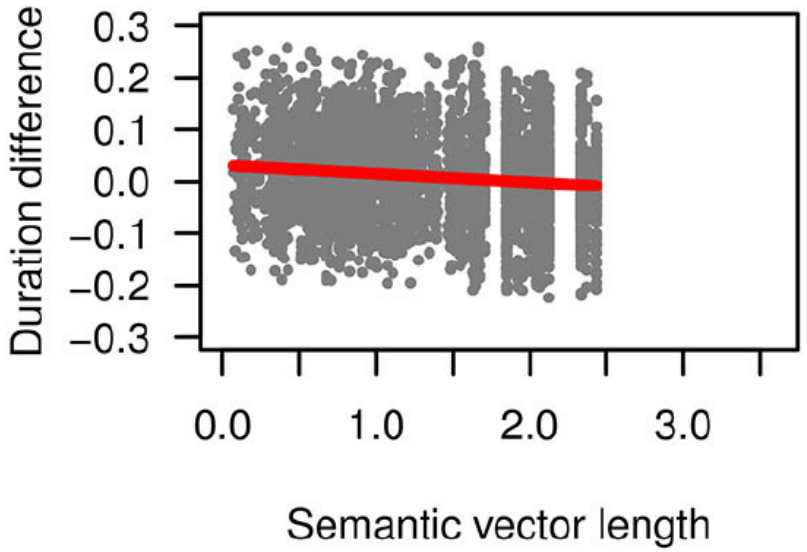
Results: Articulation/phonology





- Higher certainty in the association of form and meaning is associated with longer durations.

Results: Semantic relations



- Higher activation diversity is associated with shorter durations.

Discriminative learning: The bottom line

- Linguistic experience shapes degrees of activation which then lead to structured variability in articulatory gestures, and to different durations.

- In this way morphology can leak into what used to be called post-lexical phonology and articulation.

Summary and conclusion

- There is overwhelming evidence that morphological structure influences articulation, and thus the acoustic properties of words.
- This fact needs to be accomodated in theoretical models.
- More associative support from the lexicon leads to longer durations.
- Lexicon: network of related words, within and across paradigms
- Linguistic experience shapes degrees of association and lexical activation.
- This experience results in variability in articulatory gestures.
- In this way morphology can leak into what used to be called post-lexical phonology and articulation.
- How exactly linguistic experience translates into articulation is still unclear, but so are many non-morphological effects on phonetic implementation (e.g. word frequency)

Thank you very much for your attention!

Funding for the research presented here was provided in large parts by the

DFG Deutsche Forschungsgemeinschaft

DFG Research Unit FOR 2373
Spoken Morphology
s p ɔ ʊ k ɔ m m ɔ: f ɒ l ə d ʒ i

Grants PL 151/5-1, 151/5-2, 151/5-3, 151/7-1, 151/7-2, 151/8-1, 151/8-2

Results: Reduction and gradient paradigm uniformity

H2 The more frequent the bare stem, the shorter the bare stem (general reduction effect), and, due to paradigm uniformity, the shorter the plural stem (Gradient Paradigm Uniformity effect)

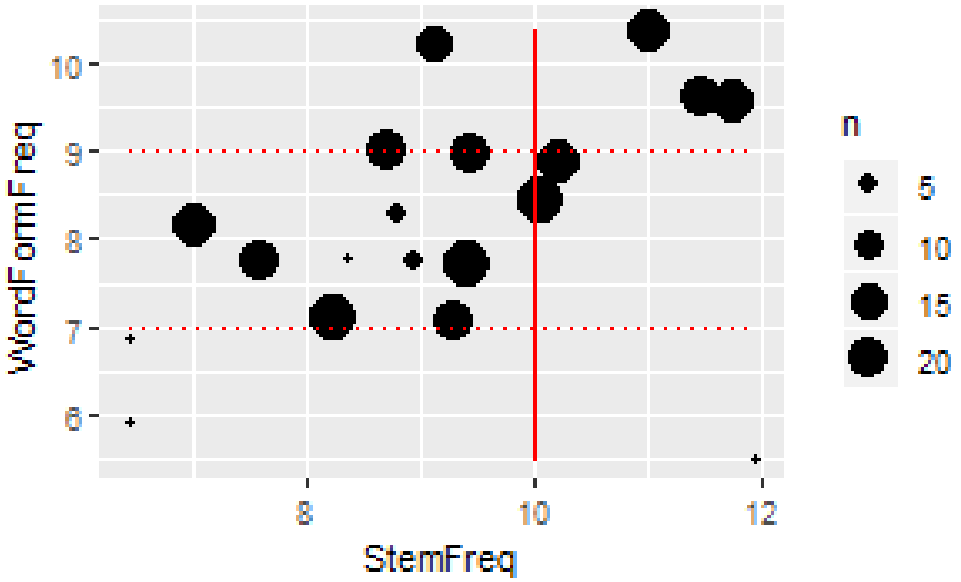
Problem: Small data set (N=295), collinearity of frequencies

Table 4: Correlation matrix for lexical frequency measures in data set 2 (rho-values, p-values are given in parentheses, Spearman test)

	WordFormFreq	RelativeFreq	StemFreq
WordFormFreq		0.14 (0.02)	0.57 (0.00)
RelativeFreq			-0.67 (0.00)

Addressing collinearity

Controlling for word-form frequency



	WordFormFreq	RelativeFreq	StemFreq
WordFormFreq		0.75 (0.000)	0.02 (0.79)
RelativeFreq			-0.60 (0.000)

Even smaller data set (N=159) ☹

Informativity hypothesis

Higher informativity leads to lengthening

- Constituent frequencies: higher frequency
  –> more expected
  –> less informative
  –> shorter
- Compound frequency: higher frequency
  –> more expected
  –> less informative
  –> shorter
- Conditional probability of N2 given N1:
  greater probability
  –> more expected
  –> less informative
  –> shorter
- Paradigmatic probability of consonants at boundary:
  greater probability
  –> more expected
  –> less informative
  –> shorter

Segmentability hypothesis

- Higher segmentability leads to lengthening
- Segmentability correlated with strength of morphological boundary
- Weaker morphological boundaries are associated with:
  - lower productivity of the category in question
  - more bound bases
  - greater semantic opacity
  - enhanced phonological integration

Segmentability hypothesis

Higher segmentability leads to lengthening

- Spelling ratio: lower spelling ratio (non-spaced / spaced)
  -> more spaced spellings
  -> more decomposable
  -> longer consonant duration at boundary
- N1 family size: greater N1 family size
  -> more productive N1
  -> more decomposable
  -> longer consonant duration at boundary

Paradigmatic support hypothesis

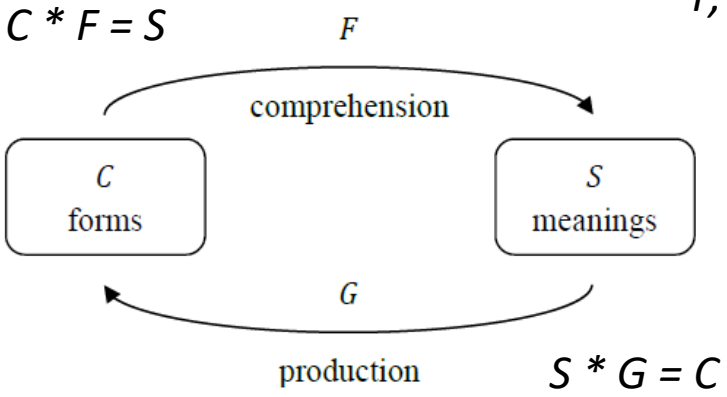
- More paradigmatic support leads to lengthening

- N1 family size: greater N1 family size
  –> more possible values for N2
  –> more spread activation/ greater uncertainty
  –> shorter consonant duration at boundary
- N1 family entropy: greater N1 entropy
  –> more spread activation/ greater uncertainty
  –> shorter consonant duration at boundary
- Paradigmatic probability of consonant after the boundary:
  higher probability
  –> less spread activation/ less uncertainty
  –> longer consonant duration

Linear Discriminative Learning

(Baayen and colleagues 2019 et seq.)

F, G = transformation matrix



Schematic example of a *C* matrix

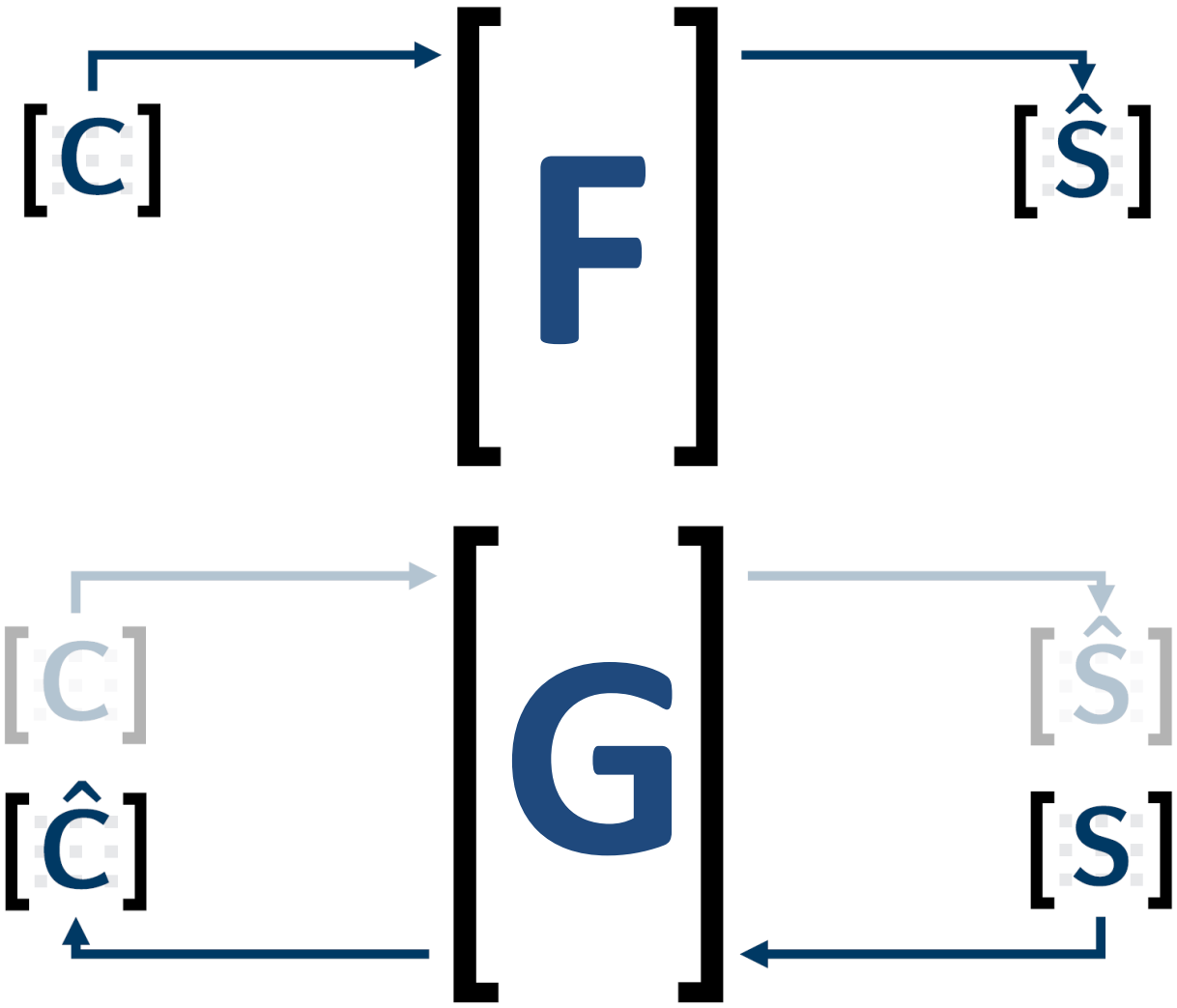
	#k{	k{t	{t#	#h{	h{p
k{t	1	1	1	0	0
h{pInIs	0	0	0	1	1
w$k	0	0	0	0	0
lEm@n	0	0	0	0	0

Schematic example of an *S* matrix

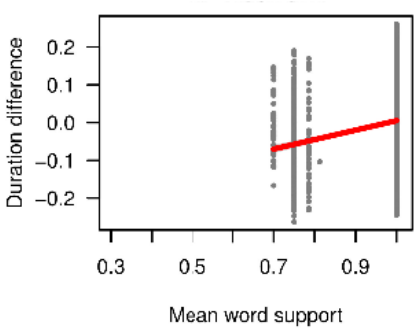
	CAT	HAPPINESS	NESS	WALK
k{t	0.000000	-6.24e-05	-0.0003179	4.71e-05
h{pInIs	-0.00056	0.0346008	0.032476	7.26e-05
w$k	0.000304	-0.0002335	-9.76e-06	0.00000
lEm@n	-7.28e-05	-2.41e-07	-0.0001247	-2.68e-05

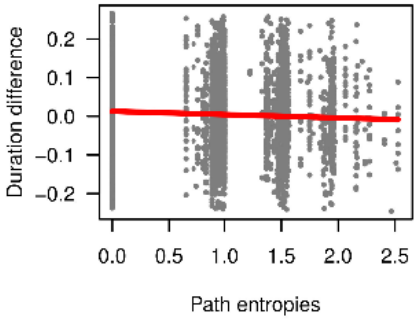
- Transformation matrices can be used to predict meanings and forms

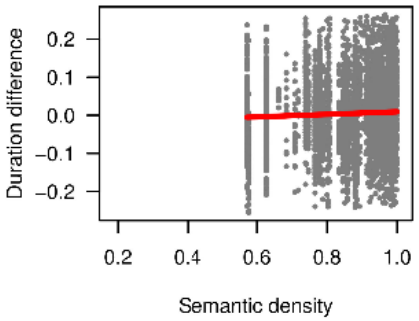
Predicting form, predicting meaning

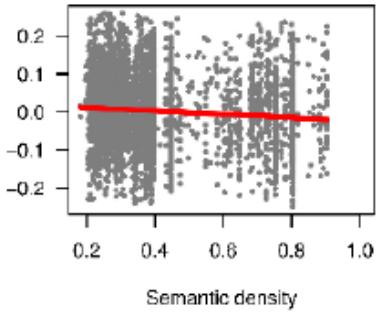


Duration in derived words: Results









predictor	represents	defined as
MEAN WORD SUPPORT	articulatory certainty	$\frac{sum\ of\ path\ supports}{number\ of\ path\ nodes}$
PATH ENTROPIES	articulatory uncertainty	Shannon entropy of path supports
SEMANTIC DENSITY	semantic transparency	mean correlation of $\hat{s}$ with top 8 neighbors

- more phonological support leads to lengthening
- more semantic support leads to both shortening and lengthening
- transparency and/or diversity?

Network setup:

word+affix only word