Simon David Stein
Ingo Plag

hhu Heinrich Heine Universität Düsseldorf
DFG Deutsche Forschungsgemeinschaft
FOR2373

Effects of morphological categories without morphemes

Modeling the acoustic duration of English derivatives with linear discriminative learning

AIMM 5

August 26–29, 2021

Background

The morpheme

Many approaches to morphology assume that complex words are concatenations of morphemes.

e.g., Selkirk 1982

- Morphemes are often understood as Saussurean “signs,” fixed units of form and meaning.
- These units are assumed to be cognitively real and represented in the lexicon.

However, the morpheme as a sign has been shown to be a theoretically problematic concept.

e.g., Matthews 1991

- Can we do without it?

What if we abandon the morpheme?

We treat all words as idiosyncratic.

We remain agnostic with regards to morphology, it's semantics all the way down.

We take words to share morphological categories.

We categorize words according to phonological and semantic similarities.

Both perspectives assume that there are no fixed units below the word level which are separately represented in the lexicon.

- How can we explore this assumption empirically?

e.g., Matthews 1991

Exploring morphological structure

Psycholinguistic evidence for morphological organization in the lexicon and in speech processing comes from speech comprehension,
but also from speech production.

- Morphological structure affects phonetic detail. — e.g., Zuraw et al. 2020, Ben Hedia 2019, Plag & Ben Hedia 2018, Hay 2007, 2003
  - e.g., no true homophony — e.g., Plag et al. 2020, Plag et al. 2017, Seyfarth et al. 2017
  - e.g., paradigmatic enhancement — e.g., Bell et al. 2020, Tomaschek et al. 2019, Cohen 2014, 2015, Kuperman et al. 2007

- Morphological information is still present at the phonetic level: Phonetic variation provides an ideal testing ground to investigate morphological structure.

Objectives

- We need to investigate whether we can model speech production with non-morphemic approaches.
- If so, we need to explore "how much" morphology we need to capture phonetic variation successfully.
  - Is it sufficient to treat all words as idiosyncratic?
  - Or should we at least build in shared morphological categories?

- What kind of model can do this?

Linear discriminative learning

- end-to-end model directly mapping forms and meanings onto each other
- associations between form and meaning are incrementally learned
- dynamic association strengths instead of fixed form-meaning units

sign-based view | discriminative view

- In LDL, morphological effects on phonetic detail can be explained by its underlying principles of learning and experience.

cf. Baayen et al. 2019b

Research questions

1. How well can LDL account for the durational variation of derivatives?
2. What do effects of LDL-derived measures tell us about speech production?
3. What does LDL tell us about the role of morphological categories?

Method

Data

		tokens	*types*	*derivational functions*
AudioBNC	audio data	4530	363	DIS, NESS, LESS, ATION, IZE

Coleman et al. 2012

Data

		tokens	*types*	*derivational functions*
AudioBNC	audio data	4530	363	DIS, NESS, LESS, ATION, IZE
AudioBNC, TASA, Baayen et al. 2019	training data		363 + 4813	DIS, NESS, LESS, ATION, IZE, AGAIN, AGENT, EE, ENCE, FUL, IC, INSTRUMENT, ISH, IST, IVE, LY, MENT, MIS, NOT, ORDINAL, OUS, OUT, SUB, UNDO, Y, MONOMORPHEMIC

Coleman et al. 2012, Ivens & Koslin 1991, Baayen et al. 2019b

Matrices

C matrix

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

S matrix

Matrices

C matrix

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

S matrix

	CAT	HAPPINESS	WALK	LEMON
k{t	0.000000	-6.24e-05	4.71e-05	-0.000138
h{plnls	-0.000110	0.0000000	0.000194	-2.20E-05
w$k	0.000304	-0.0002335	0.000000	-3.74E-05
lEm@n	-7.28e-05	-2.41e-07	-2.68e-05	0.00000

Matrices

learning algorithm in TASA
Baayen et al. 2019

752,130 sentences,
10,719,386 tokens

lexome-to-lexome matrix

	CAT	HAPPINESS	NESS	WALK
CAT	0.000000	-6.24E-05	-0.0003179	4.71E-05
HAPPINESS	-0.000110	0.00000000	0.032476	0.000194
NESS	-0.000450	0.0346008	0.000000	-0.0001
WALK	0.000304	-0.0002335	-9.76E-06	0.000000
LEMON	-7.28E-05	-2.41E-07	-0.0001247	-2.68E-05

C matrix

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

S matrix

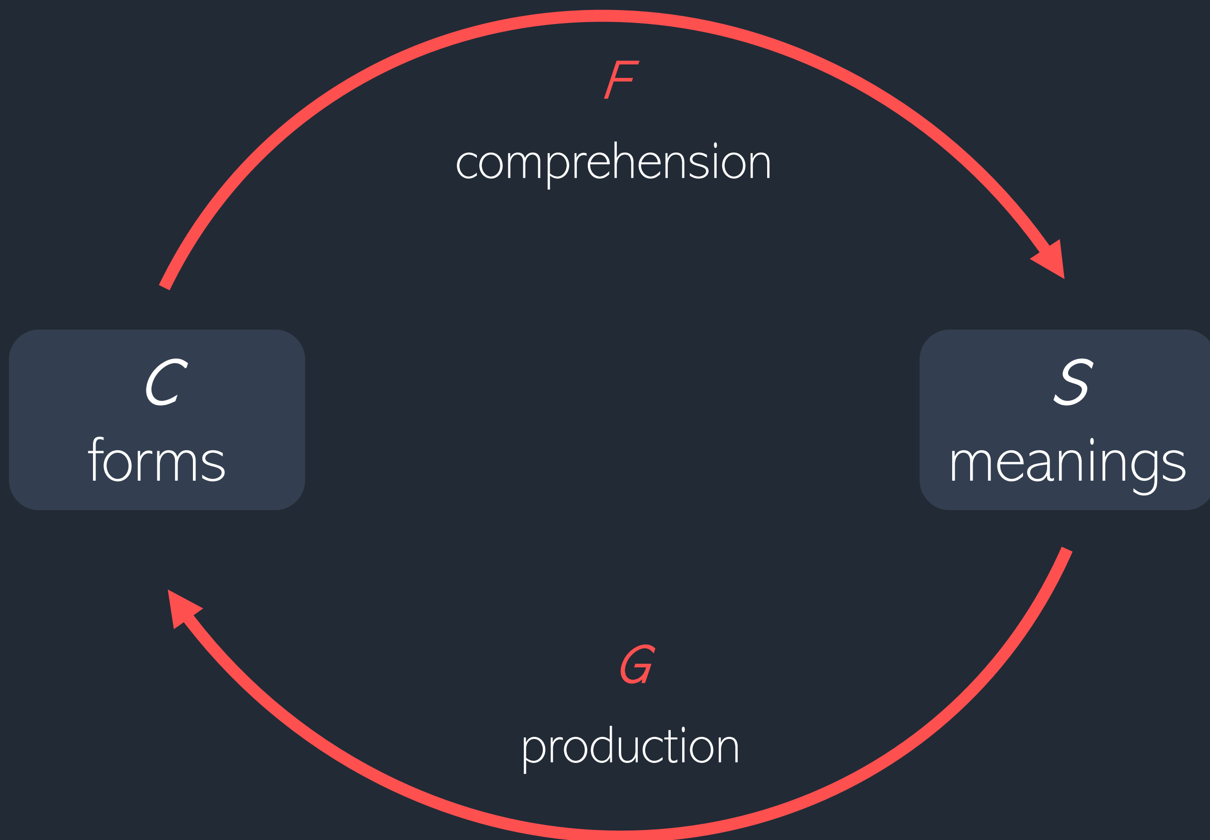
	CAT	HAPPINESS	WALK	LEMON
k{t	0.000000	-6.24e-05	4.71e-05	-0.000138
h{plnls	-0.000110	0.0000000	0.000194	-2.20E-05
w$k	0.000304	-0.0002335	0.000000	-3.74E-05
lEm@n	-7.28e-05	-2.41e-07	-2.68e-05	0.00000

Baayen et al. 2019b

Two networks

Idiosyncratic Network	Morphology Network
$\overrightarrow{happiness}$	$\overrightarrow{happiness} + \overrightarrow{NESS}$
Vectors do not contain explicit information about morphological function	Vectors do contain explicit information about morphological function

Comprehension and production mapping



Comprehension and production mapping

predicting meanings

$$\hat{S} = CF$$

predicting forms

$$\hat{C} = SG$$

Modeling durations

linear models and mixed effects models with random intercept for word type

response variable

- DURATION DIFFERENCE
  residuals of a linear model OBSERVED DURATION ~ BASELINE DURATION

predictors

- MEAN WORD SUPPORT
- PATH ENTROPIES
- SEMANTIC VECTOR LENGTH
- SEMANTIC DENSITY
- TARGET CORRELATION
- SPEECH RATE

Results

Network accuracy

	Idiosyncratic Network	Morphology Network
comprehension	81 %	82 %
production	99 %	99 %

Explained variance of variables predicting duration

	Idiosyncratic Network	Morphology Network
R^2 adj. lm	.38	.37
R^2 mar. lmer	.37	.36

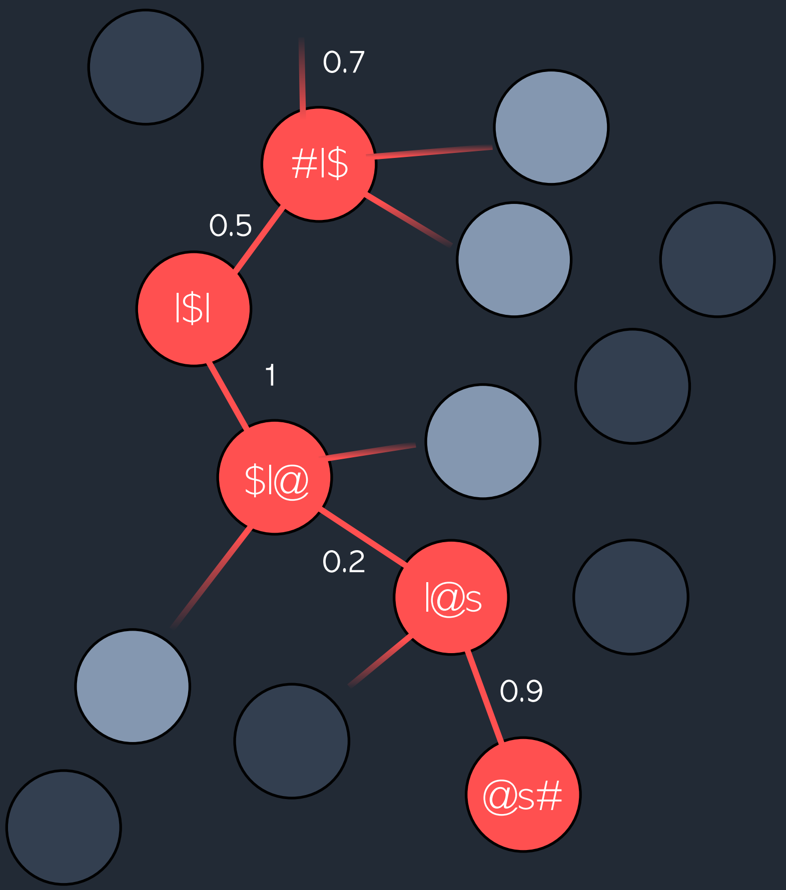
traditional model with
WORD FREQUENCY, RELATIVE FREQUENCY, BIGRAM FREQUENCY, BIPHONE PROBABILITY, AFFIX, SPEECH RATE

R^2 adj. lm	.37
R^2 mar. lmer	.37

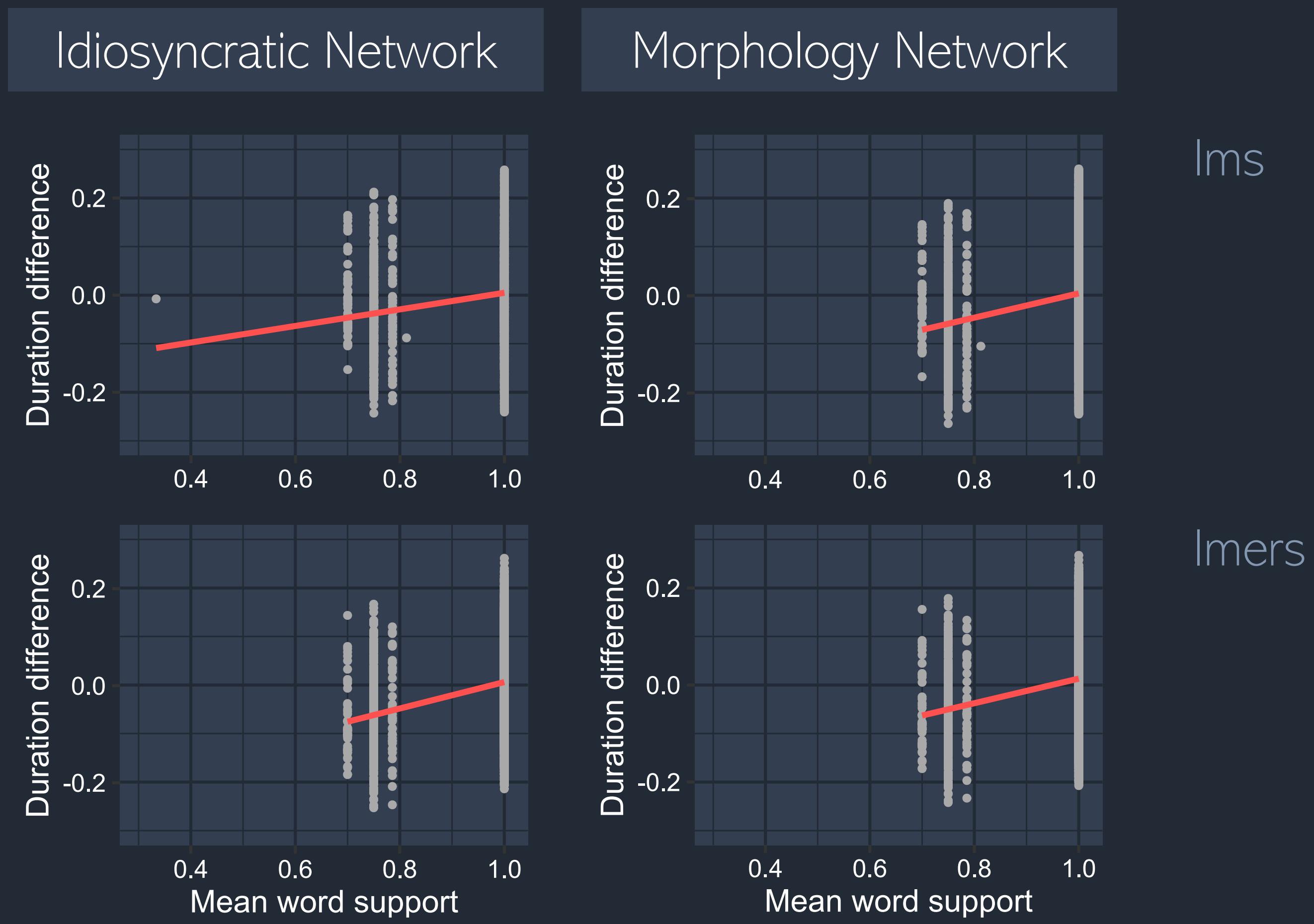
MEAN WORD SUPPORT

$$\frac{\textit{sum of path supports}}{\textit{number of path nodes}}$$

can represent
articulatory certainty



MEAN WORD SUPPORT

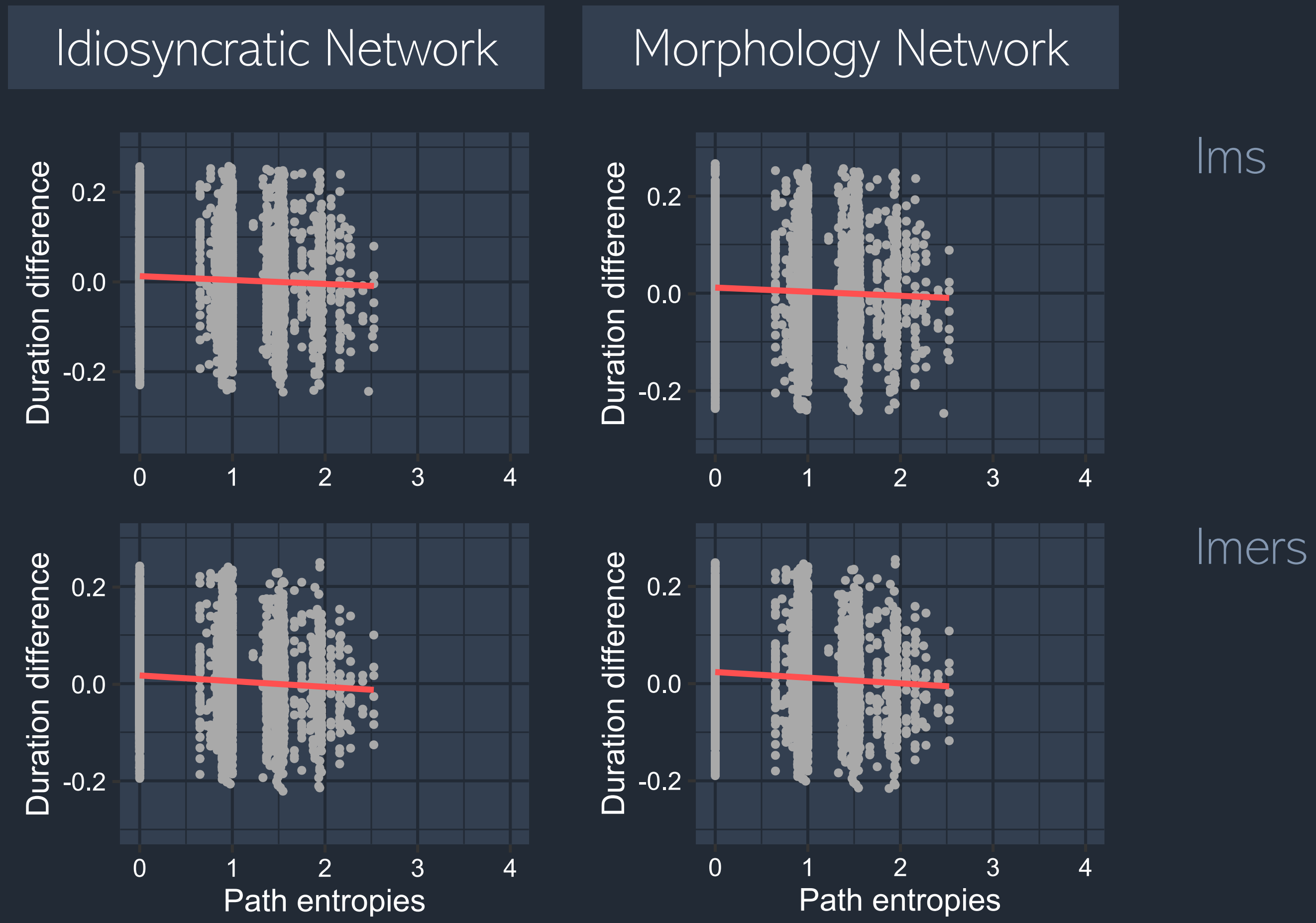


PATH ENTROPIES

Shannon entropy
of path supports

can represent
articulatory uncertainty

PATH ENTROPIES



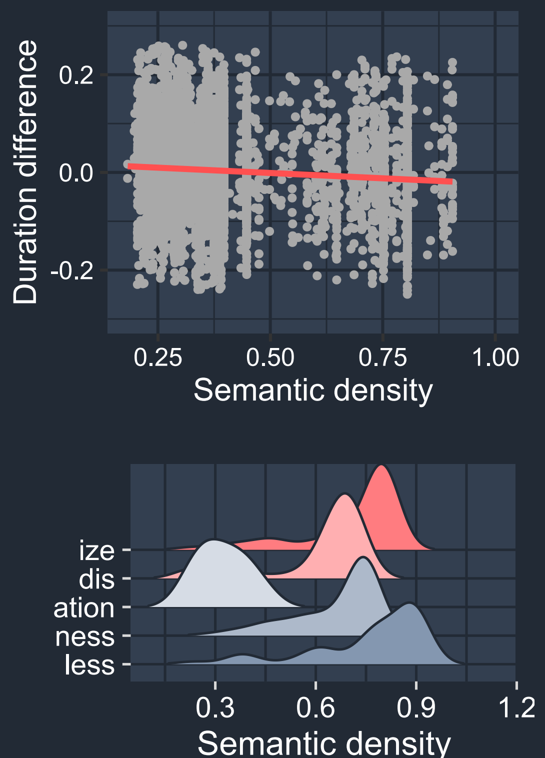
SEMANTIC DENSITY

mean correlation of $\hat{s}$
with top 8 neighbors

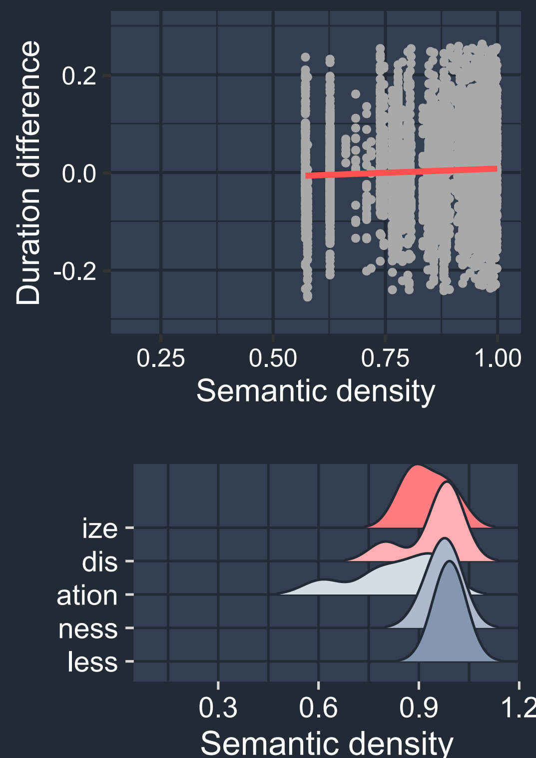
can represent
semantic transparency

SEMANTIC DENSITY

Idiosyncratic Network



Morphology Network



lms

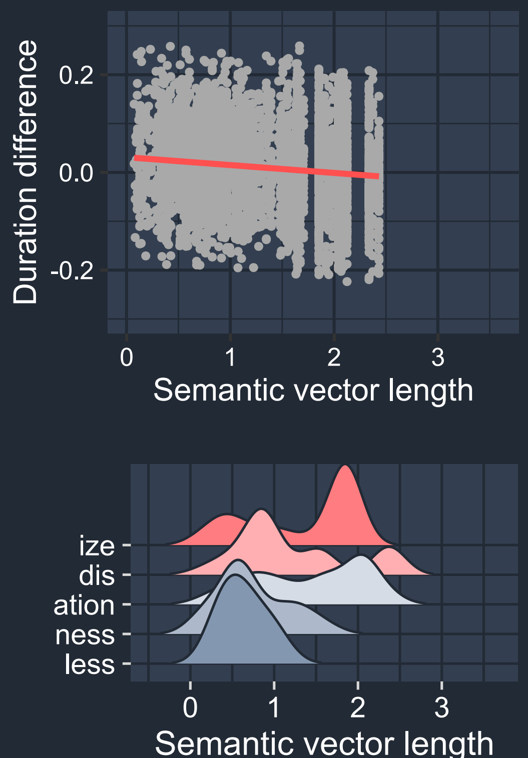
SEMANTIC VECTOR LENGTH

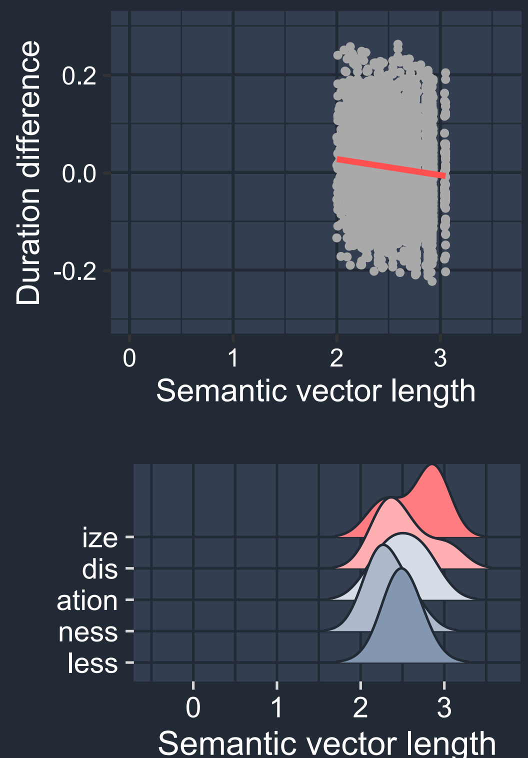
L1 distance of $\hat{s}$

can represent
activation diversity or polysemy

SEMANTIC VECTOR LENGTH

Idiosyncratic Network	Morphology Network





lmers

Discussion

1. How well can LDL account for the durational variation of derivatives?

LDL-derived variables are successful in predicting derivative durations.

- This is further evidence that discriminative models are a promising approach to speech production.
- We can model morpho-phonetic effects without referring to morphemic structure.

cf., e.g., Baayen et al. 2019, Chuang et al. 2020, Tomaschek et al. 2019, Tucker et al. 2019

2. What do effects of LDL-derived measures tell us about speech production?

Longer or shorter durations are predicted by various measures relating to different notions, such as:

- certainty
- semantic transparency
- activation diversity

These findings are generally in line with the literature.

Zuraw et al. 2020,
Tucker et al. 2019,
Tomaschek et al. 2019,
Plag & Ben Hedia 2018,
Cohen 2015,
Cohen 2014,
Plag & Baayen 2009,
Hay 2007,
Kuperman et al. 2007,
Hay 2003,
Schreuder & Baayen 1997

3. What does LDL tell us about the role of morphological categories?

Differences between morphological categories emerge from:

- the network with explicit vectors for morphological functions, and
- the network without explicit vectors for morphological functions.

- We don't need to hard-code morphology into our word representations.

3. What does LDL tell us about the role of morphological categories?

In addition, some of the differences between morphological categories mirror traditional classifications from the literature.

- Semantic density is higher for words with NESS, LESS and DIS than for words with ATION (cf. transparency of *-ness*, *-less*, and *dis-* vs. *-ation*).
- Semantic vector length is highest for IZE and ATION words (cf. semantics of *-ize* and *-ation* vs. *-less*, *dis-*, and *-ness*).

cf. Bauer et al. 2013; Plag 2018

Conclusion

Takeaways

- Differences between morphological categories are successfully captured by the semantic vectors in both networks.
- Morpho-phonetic effects can be modeled without reference to the morpheme.

Thank you!

- Arnold, D., Tomaschek, F., Sering, K., Lopez, F., and Baayen, R. H. (2017). Words from spontaneous conversational speech can be recognized with human-like accuracy by an error-driven learning algorithm that discriminates between meanings straight from smart acoustic features, bypassing the phoneme as recognition unit. *PLoS ONE* 12, e0174623. doi: 10.1371/journal.pone.0174623.
- Baayen, R. H. (2008). *Analyzing Linguistic Data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Baayen, R. H., Chuang, Y.-Y., and Heitmeier, M. (2019a). *WpmWithLdl: Implementation of Word and Paradigm Morphology with Linear Discriminative Learning*. R package. http://www.sfs.uni-tuebingen.de/~hbaayen/software.html.
- Baayen, R. H., Chuang, Y.-Y., Shafaei-Bajestan, E., and Blevins, J. P. (2019b). 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.
- Baayen, R. H., and Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research* 3, 12–28. doi: 10.21500/20112084.807.
- Baayen, R. H., Milin, P., **Đurđević**, D. F., Hendrix, P., and Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review* 118, 438–481. doi: 10.1037/a0023851.
- Baayen, R. H., Milin, P., and Ramscar, M. (2016). Frequency in lexical processing. *Aphasiology* 30, 1174–1220.

- Baayen, R. H., Piepenbrock, R., and Gulikers, L. (1995). *The CELEX Lexical Database*. Philadelphia: Linguistic Data Consortium.
- Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67, 1–48. doi: 10.18637/jss.v067.i01.
- Bauer, L., Lieber, R., and Plag, I. (2013). *The Oxford reference guide to English morphology*. Oxford: Oxford University Press.
- Bell, A., Brenier, J. M., Gregory, M., Girand, C., and Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language* 60, 92–111.
- Ben Hedia, S. (2019). *Gemination and degemination in English affixation: Investigating the interplay between morphology, phonology and phonetics*. Berlin: Language Science Press.
- Ben Hedia, S., and Plag, I. (2017). Gemination and degemination in English prefixation. Phonetic evidence for morphological organization. *Journal of Phonetics* 62, 34–49.
- Beręsewicz, M. (2015). calc.relip.mm: Variable importance for mixed models. https://gist.github.com/BERENZ/e9b581a4b7160357934e.
- Bertram, R., Baayen, R. H., and Schreuder, R. (2000). Effects of family size for complex words. *Journal of Memory and Language* 42, 390–405. doi: 10.1006/jmla.1999.2681.
- Blevins, J. P. (2016). The minimal sign. In *The Cambridge handbook of morphology*, ed. A. Hippisley and G. Stump (Cambridge: Cambridge University Press), 50–69.
- Boersma, P., and Weenik, D. J. M. (2001). *Praat: Doing phonetics by computer*. http://www.praat.org/.
- Burnage, G. (1990). *CELEX: A guide for users*. Nijmegen: Centre for Lexical Information.

- Caselli, N. K., Caselli, M. K., and Cohen-Goldberg, A. M. (2016). Inflected words in production. Evidence for a morphologically rich lexicon. *The Quarterly Journal of Experimental Psychology* 69, 432–454.
- Chatterjee, S., and Hadi, A. S. (2006). *Regression analysis by example*. 4th ed. Hoboken: John Wiley & Sons.
- Chomsky, N., and Halle, M. (1968). *The sound pattern of English*. New York, Evanston, London: Harper and Row.
- Chuang, Y.-Y., Vollmer, M. L., Shafaei-Bajestan, E., Gahl, S., Hendrix, P., and Baayen, R. H. (2020). The processing of pseudoword form and meaning in production and comprehension. A computational modeling approach using linear discriminative learning. *Behavior Research Methods*. doi: 10.3758/s13428-020-01356-w.
- Cohen, C. (2014). Probabilistic reduction and probabilistic enhancement. Contextual and paradigmatic effects on morpheme pronunciation. *Morphology* 24, 291–323.
- Cohen, C. (2015). Context and paradigms. *The Mental Lexicon* 10, 313–338. doi: 10.1075/ml.10.3.01coh.
- Coleman, J., Baghai-Ravary, L., Pybus, J., and Grau, S. (2012). *Audio BNC: The audio edition of the Spoken British National Corpus*. http://www.phon.ox.ac.uk/AudioBNC.
- Davies, M. (2008). *The Corpus of Contemporary American English: 450 million words, 1990–present*. http://corpus.byu.edu/coca/.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review* 93, 283–321. doi: 10.1037/0033-295X.93.3.283.

- Divjak, D. (2019). *Frequency in language: Memory, attention and learning. Cambridge*: Cambridge University Press.
- Edwards, J., Beckman, M. E., and Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research* 47, 421–436. doi: 10.1044/1092-4388(2004/034).
- Engemann, U. M., and Plag, I. (2021). Phonetic reduction and paradigm uniformity effects in spontaneous speech. *The Mental Lexicon* 16, 165–198.
- Filipović Đurđević, D., and Kostić, A. (2021). We probably sense sense probabilities. *Language, Cognition and Neuroscience*, 1–28. doi: 10.1080/23273798.2021.1909083.
- Firth, J. R. (1957). A synopsis of linguistic theory, 1930–1955. In *Studies in linguistic analysis* (Oxford: Blackwell), 1–31.
- Gahl, S., Yao, Y., and Johnson, K. (2012). Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *Journal of Memory and Language* 66, 789–806. doi: 10.1016/j.jml.2011.11.006.
- Grömping, U. (2006). Relative Importance for Linear Regression in R. The Package relaimpo. *Journal of Statistical Software* 17. doi: 10.18637/jss.v017.i01.
- Hay, J. (2001). Lexical frequency in morphology. Is everything relative? *Linguistics* 39, 1041–1070.
- Hay, J. (2003). Causes and consequences of word structure. New York, London: Routledge.
- Hay, J. (2007). The phonetics of un. In *Lexical creativity, texts and contexts*, ed. J. Munat (Amsterdam, Philadelphia: John Benjamins), 39–57.

- Howes, D. H., and Solomon, R. L. (1951). Visual duration threshold as a function of word-probability. *Journal of Experimental Psychology* 41, 401–410.
- Ivens, S. H., and Koslin, B. L. (1991). *Demands for reading literacy require new accountability measures*. Brewster: Touchstone Applied Science Associates.
- Kiparsky, P. (1982). Lexical morphology and phonology. In *Linguistics in the morning calm: Selected papers from SICOL*, ed. I.-S. Yang (Seoul: Hanshin), 3–91.
- Kunter, G. (2016). *Coquery: A free corpus query tool*. www.coquery.org.
- Kuperman, V., Pluymaekers, M., Ernestus, M., and Baayen, R. H. (2007). Morphological predictability and acoustic duration of interfixes in Dutch compounds. *The Journal of the Acoustical Society of America* 121, 2261–2271. doi: 10.1121/1.2537393.
- Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2016). lmerTest: Tests in linear fixed effects models. R package. https://cran.r-project.org/web/packages/lmerTest/index.html.
- Ladefoged, P., and Johnson, K. (2011). *A course in phonetics*. 6th ed. Boston: Wadsworth Cengage Learning.
- Levelt, W. J. M., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22, 1–38.
- Lindeman, R. H., Merenda, P. F., and Gold, R. Z. (1980). *Introduction to bivariate and multivariate analysis*. Glenview, London: Scott, Foresman and Company.
- Machač, P., and Skarnitzl, R. (2009). *Principles of phonetic segmentation*. Prague: Epocha Publishing House.
- Matthews, Peter H. 1991. *Morphology: An introduction to the theory of word structure*, 2nd edn. (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

- Moore, E. H. (1920). On the reciprocal of the general algebraic matrix. *Bulletin of the American Mathematical Society* 26, 394–395.
- Munson, B. (2001). Phonological pattern frequency and speech production in adults and children. *Journal of Speech, Language, and Hearing Research* 44, 778–792. doi: 10.1044/1092-4388(2001/061).
- Penrose, R. (1955). A generalized inverse for matrices. *Mathematical Proceedings of the Cambridge Philosophical Society 51*, 406–413. doi: 10.1017/S0305004100030401.
- Pirrelli, V., Marzi, C., Ferro, M., Cardillo, F. A., Baayen, R. H., and Milin, P. (2020). Psycho-computational modelling of the mental lexicon: A discriminative learning perspective. In *Word knowledge and word usage: A cross-disciplinary guide to the mental lexicon*, ed. V. Pirrelli, I. Plag, and W. U. Dressler (Berlin, Boston: Mouton de Gruyter), 23–82.
- Plag, I. (2018). *Word-formation in English*. 2nd ed. Cambridge: Cambridge University Press.
- Plag, I., and Baayen, R. H. (2009). Suffix ordering and morphological processing. *Language* 85, 109–152. doi: 10.1353/lan.0.0087.
- Plag, I., and Balling, L. W. (2020). Derivational morphology: An integrative perspective on some fundamental questions. In *Word knowledge and word usage: A cross-disciplinary guide to the mental lexicon*, ed. V. Pirrelli, I. Plag, and W. U. Dressler (Berlin, Boston: Mouton de Gruyter), 295–335.
- Plag, I., and Ben Hedia, S. (2018). The phonetics of newly derived words: Testing the effect of morphological segmentability on affix duration. In *Expanding the lexicon: Linguistic innovation, morphological productivity, and ludicity*, ed. S. Arndt-Lappe, A. Braun, C. Moulin, and E. Winter-Froemel (Berlin, New York: Mouton de Gruyter), 93–116.

- Plag, I., Dalton-Puffer, C., and Baayen, R. H. (1999). Morphological productivity across speech and writing. *English Language and Linguistics* 3, 209–228. doi: 10.1017/S1360674399000222.
- Plag, I., Homann, J., and Kunter, G. (2017). Homophony and morphology. The acoustics of word-final S in English. *Journal of Linguistics* 53, 181–216.
- Plag, I., Lohmann, A., Ben Hedia, S., and Zimmermann, J. (2020). An <s> is an <s'>, or is it? Plural and genitive-plural are not homophonous. In *Complex words: Advances in morphology*, ed. L. Körtvélyessy and P. Stekauer (Cambridge: Cambridge University Press).
- Pluymaekers, M., Ernestus, M., and Baayen, R. H. (2005a). Articulatory planning is continuous and sensitive to informational redundancy. *Phonetica* 62, 146–159.
- Pluymaekers, M., Ernestus, M., and Baayen, R. H. (2005b). Lexical frequency and acoustic reduction in spoken Dutch. *The Journal of the Acoustical Society of America* 118, 2561–2569.
- R Core Team (2020). *R: A language and environment for statistical computing*. http://www.R-project.org/.
- Ramscar, M., and Yarlett, D. (2007). Linguistic self-correction in the absence of feedback. A new approach to the logical problem of language acquisition. *Cognitive Science* 31, 927–960. doi: 10.1080/03640210701703576.
- Ramscar, M., Yarlett, D., Dye, M., Denny, K., and Thorpe, K. (2010). The effects of feature-label-order and their implications for symbolic learning. *Cognitive Science* 34, 909–957. doi: 10.1111/j.1551-6709.2009.01092.x.
- Roelofs, A., and Ferreira, V. S. (2019). The architecture of speaking. In *Human language: From genes and brains to behavior*, ed. P. Hagoort (Cambridge, Massachusetts: MIT Press), 35–50.

- Saussure, F. de (1916). *Cours de linguistique générale*. Paris: Payot.
- Schreuder, R., and Baayen, R. H. (1997). How complex simplex words can be. *Journal of Memory and Language* 37, 118–139.
- Schuppler, B., van Dommelen, W. A., Koreman, J., and Ernestus, M. (2012). How linguistic and probabilistic properties of a word affect the realization of its final /t/. Studies at the phonemic and sub-phonemic level. *Journal of Phonetics* 40, 595–607.
- Selkirk, Elisabeth O. 1982. *The syntax of words* (Linguistic Inquiry Monographs 7). Cambridge, Mass., London: MIT Press.
- Seyfarth, S., Garellek, M., Gillingham, G., Ackerman, F., and Malouf, R. (2017). Acoustic differences in morphologically-distinct homophones. *Language, Cognition and Neuroscience* 33, 32–49.
- Shafaei-Bajestan, E., Moradipour-Tari, M., Uhrig, P., and Baayen, R. H. (2020). LDL-AURIS: Error-driven learning in modeling spoken word recognition. *PsyArXiv*, 1–36. doi: 10.31234/osf.io/v6cu4.
- Sóskuthy, M., and Hay, J. (2017). Changing word usage predicts changing word durations in New Zealand English. *Cognition* 166, 298–313.
- Tomaschek, F., Plag, I., Ernestus, M., and Baayen, R. H. (2019). Phonetic effects of morphology and context. Modeling the duration of word-final S in English with naïve discriminative learning. *Journal of Linguistics*, 1–39.
- Torreira, F., and Ernestus, M. (2009). Probabilistic effects on French [t] duration. *INTERSPEECH*, 448–451.
- Tucker, B., Sims, M., and Baayen, R. H. (2019). Opposing forces on acoustic duration. Preprint submitted to Elsevier. *PsyArXiv*, 1–38. doi: 10.31234/osf.io/jc97w.

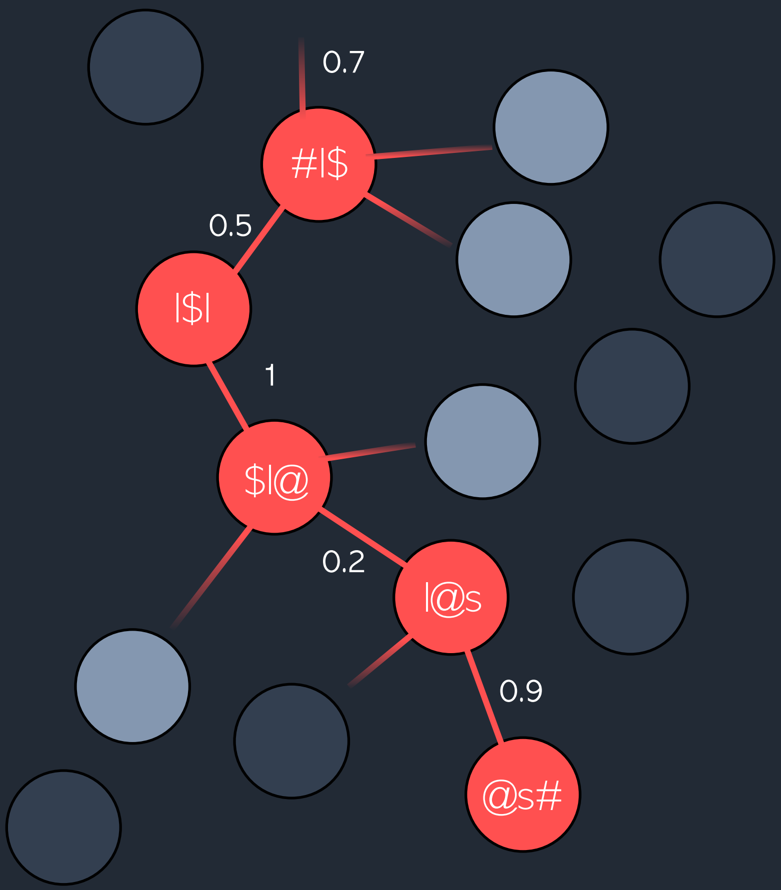
- Tucker, B. V., and Ernestus, M. (2016). Why we need to investigate casual speech to truly understand language production, processing and the mental lexicon. *The Mental Lexicon* 11, 375–400. doi: 10.1075/ml.11.3.03tuc.
- Turk, A., and Shattuck-Hufnagel, S. (2020). *Speech timing: Implications for theories of phonology, speech production, and speech motor control*. New York: Oxford University Press.
- Turnbull, R. (2018). Patterns of probabilistic segment deletion/reduction in English and Japanese. *Linguistics Vanguard* 4. doi: 10.1515/lingvan-2017-0033.
- Vitevitch, M. S., and Luce, P. A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, and Computers* 36, 481–487.
- Walsh, L., Hay, J., Derek, B., Grant, L., King, J., Millar, P., Papp, V., and Watson, K. (2013). The UC QuakeBox Project. Creation of a community-focused research archive. *New Zealand English Journal* 27, 20–32.
- Widrow, B., and Hoff, M. E. (1960). Adaptive switching circuits. *WESCON Convention Record Part IV*, 96–104.
- Zimmerer, F., Scharinger, M., and Reetz, H. (2014). Phonological and morphological constraints on German /t/-deletions. *Journal of Phonetics* 45, 64–75. doi: 10.1016/j.wocn.2014.03.006.
- Zuraw, K., Lin, I., Yang, M., and Peperkamp, S. (2020). Competition between whole-word and decomposed representations of English prefixed words. *Morphology*, 10.1007/s11525-020-09354-6. doi: 10.1007/s11525-020-09354-6.

Appendix

MEAN WORD SUPPORT

$$\frac{\textit{sum of path supports}}{\textit{number of path nodes}}$$

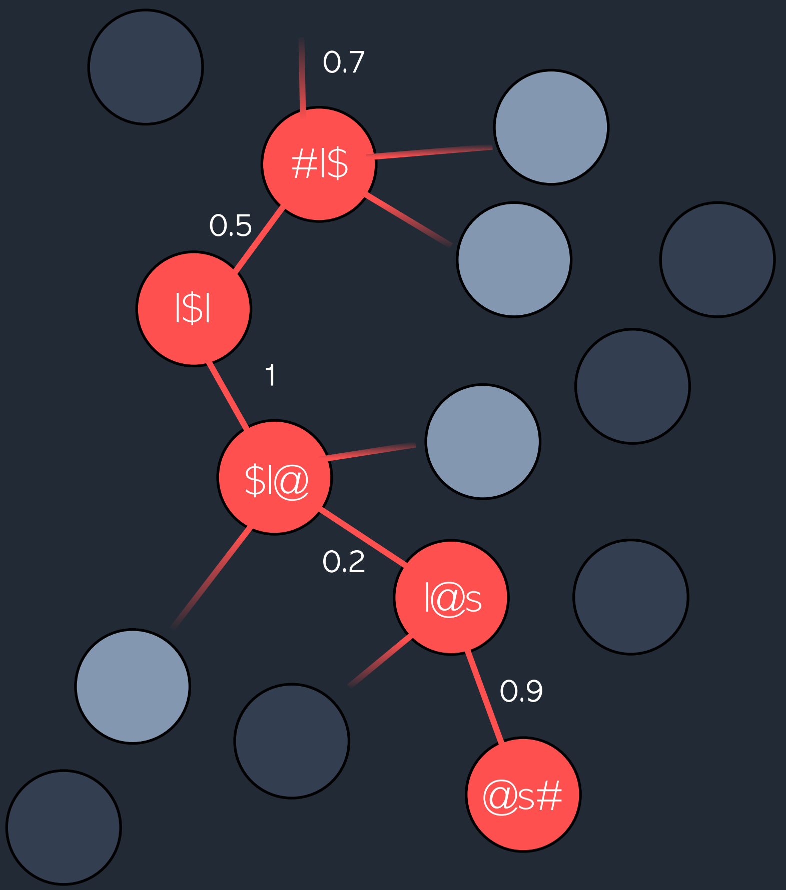
can represent
articulatory certainty



PATH ENTROPIES

Shannon entropy
of path supports

can represent
articulatory uncertainty



Three networks

Idiosyncratic Network	Morphology Network	Base Network
$\overrightarrow{happiness}$	$\overrightarrow{happiness} + \overrightarrow{NESS}$	$\overrightarrow{happy} + \overrightarrow{NESS}$
Vectors contain:	Vectors contain:	Vectors contain:
idiosyncratic information about derivative,	idiosyncratic information about derivative,	no idiosyncratic information about derivative,
no information about morphological function	information about morphological function	information about morphological function

Network accuracy

	Idiosyncratic Network	Morphology Network	Base Network
comprehension	81 %	82 %	83 %
production	99 %	99 %	98 %

Similarity of semantic matrices

Idiosyncratic Network ⟷ Morphology Network	r = .08
Idiosyncratic Network ⟷ Base Network	r = .1
Base Network ⟷ Morphology Network	r = .9

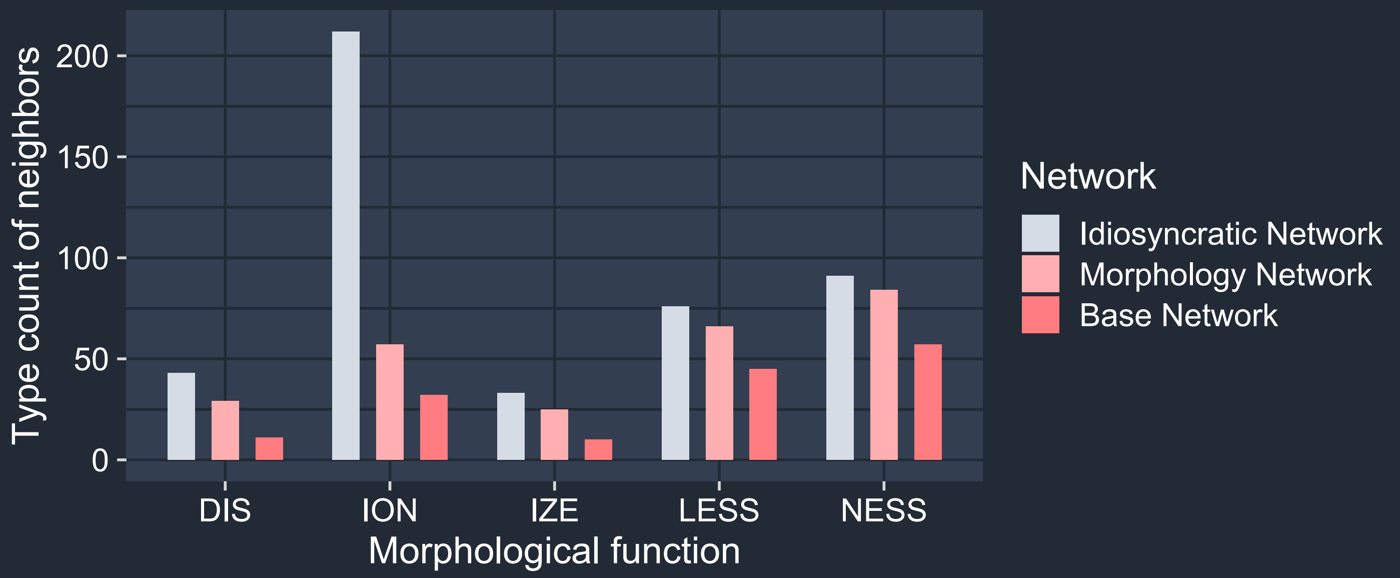
Explained variance of variables predicting duration

	Idiosyncratic Network	Morphology Network	Base Network
R^2 adj. lm	.38	.37	.36
R^2 mar. lmer	.37	.36	.35

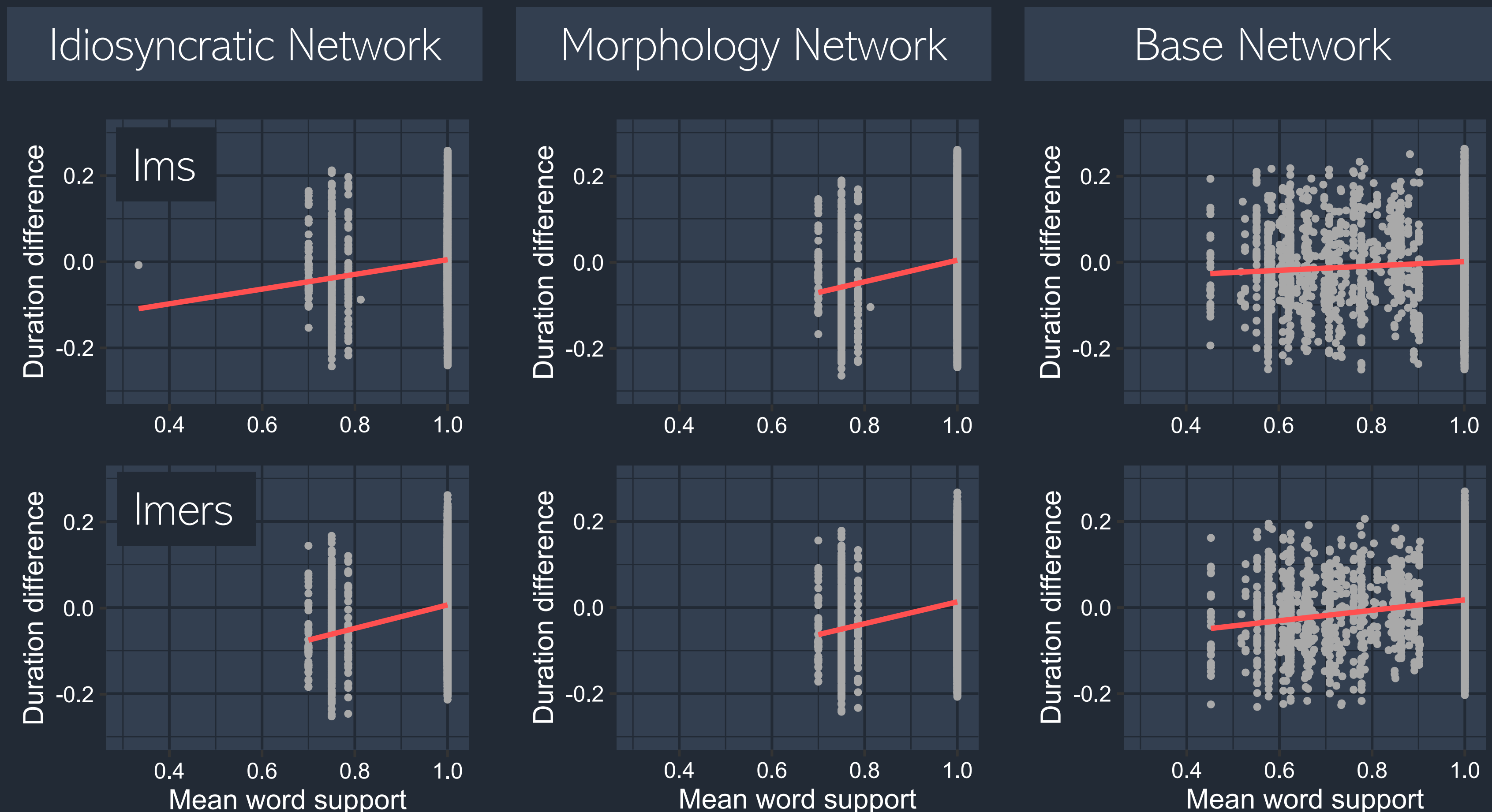
traditional model with
WORD FREQUENCY, RELATIVE FREQUENCY, BIGRAM FREQUENCY, BIPHONE PROBABILITY, AFFIX, SPEECH RATE

R^2 adj. lm	.37
R^2 mar. lmer	.37

Type count of top 8 neighbors

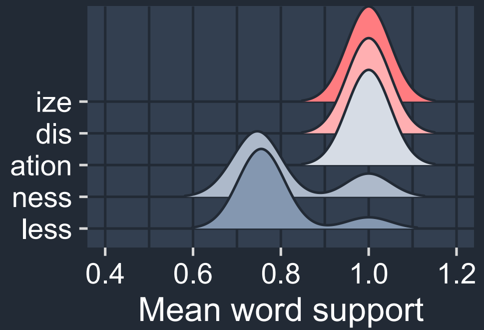


MEAN WORD SUPPORT

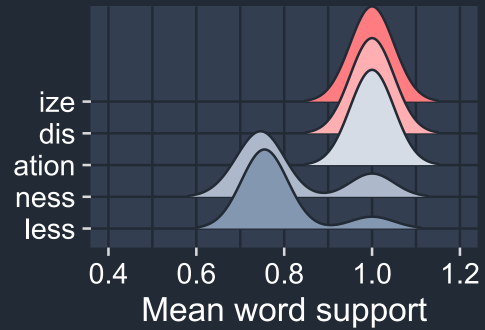


MEAN WORD SUPPORT

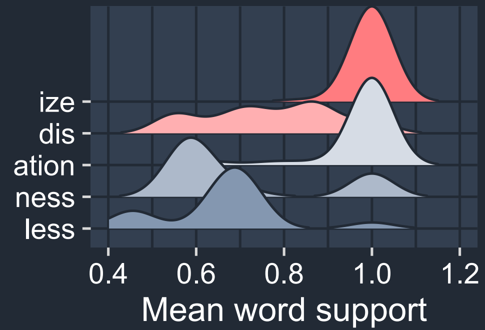
Idiosyncratic Network



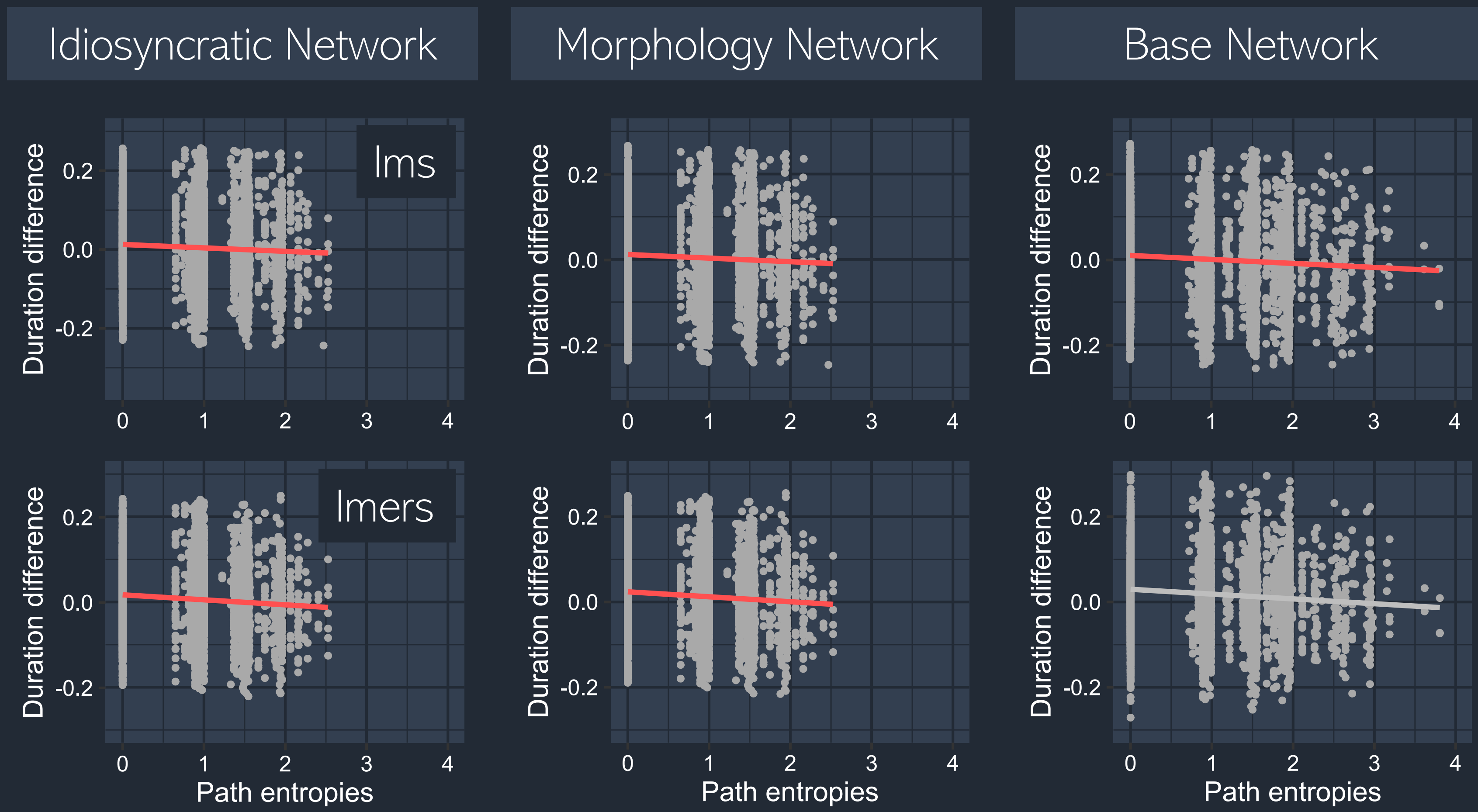
Morphology Network



Base Network

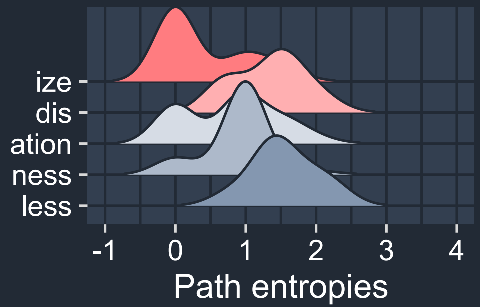


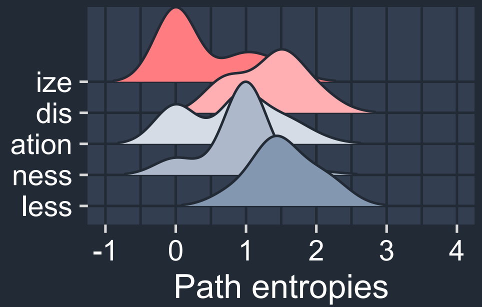
PATH ENTROPIES

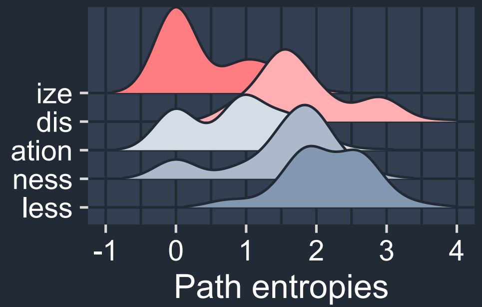


PATH ENTROPIES

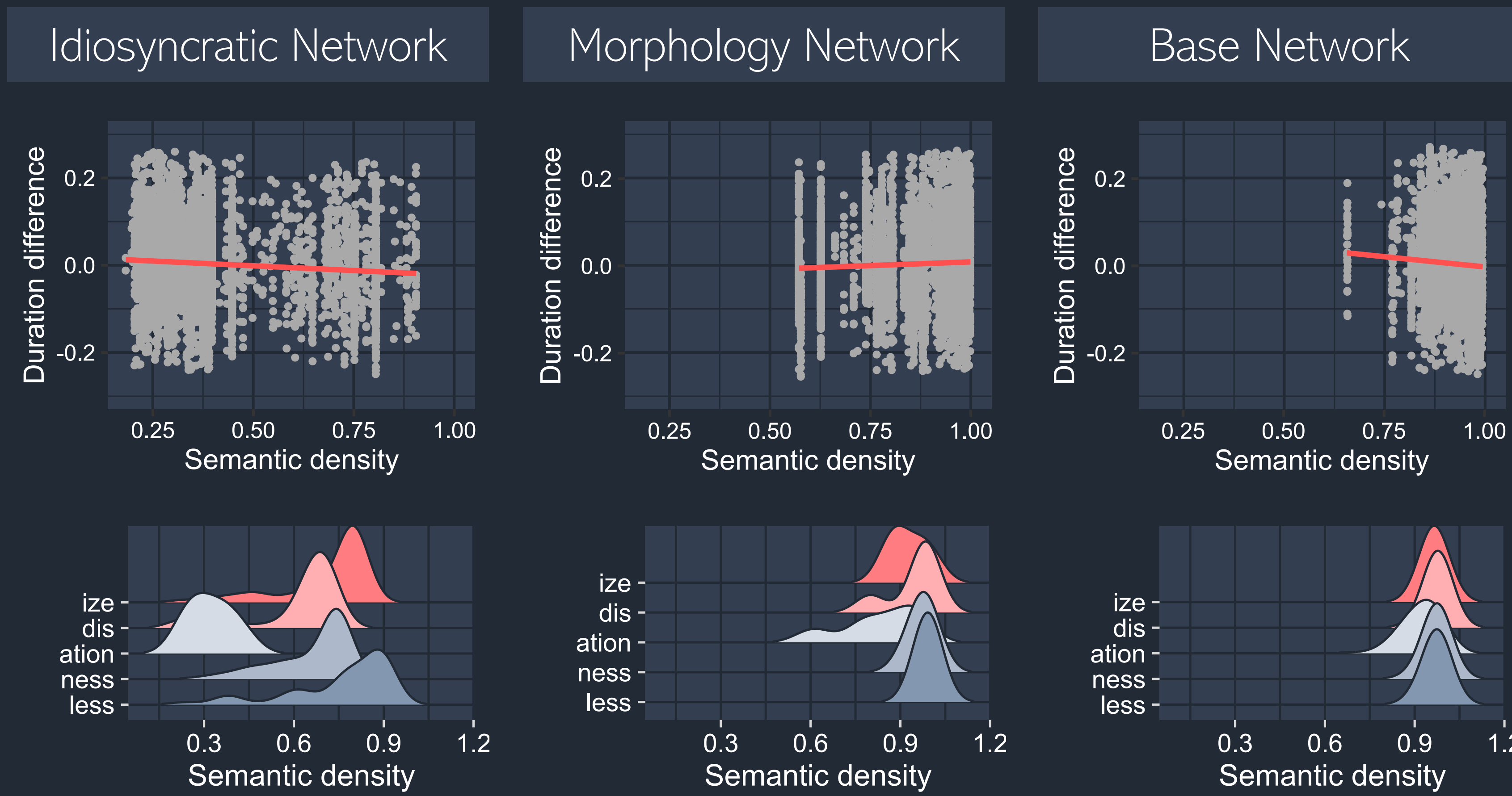
Idiosyncratic Network	Morphology Network	Base Network





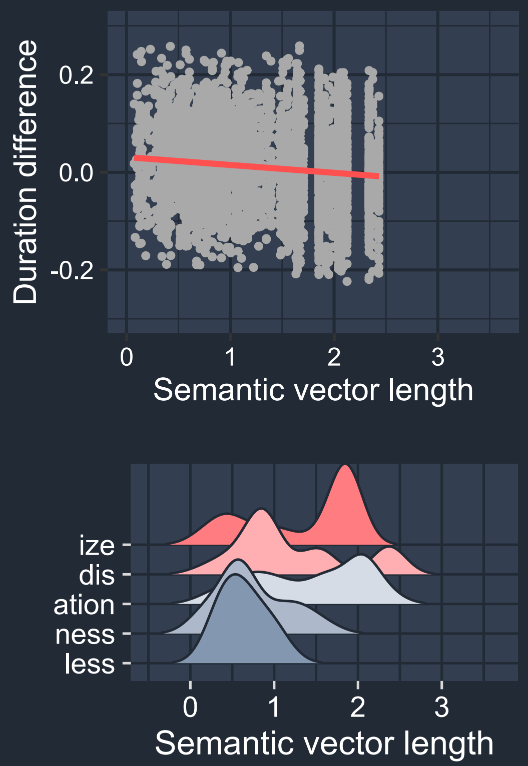


SEMANTIC DENSITY lms

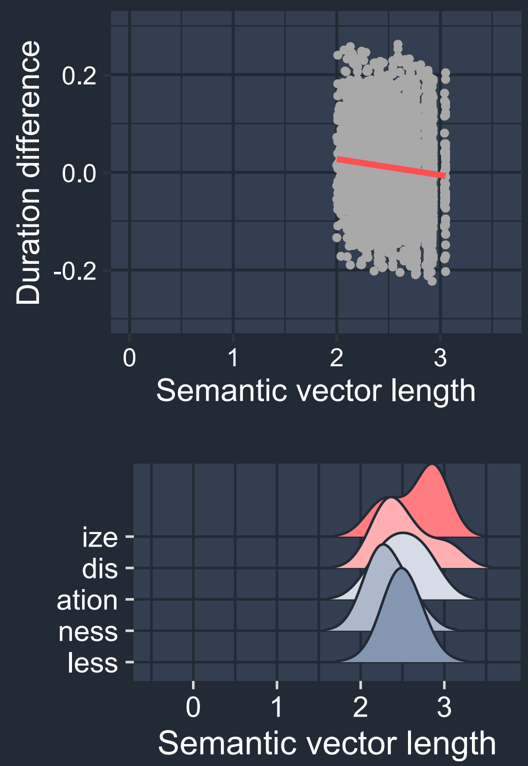


SEMANTIC VECTOR LENGTH lmers

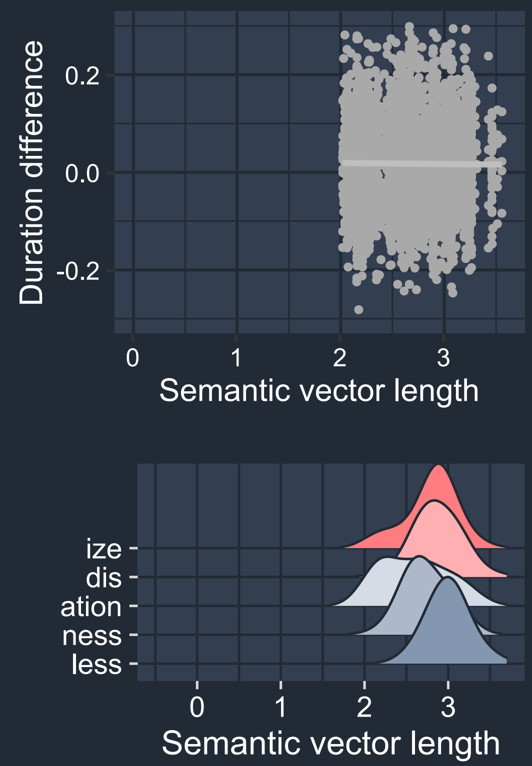
Idiosyncratic Network



Morphology Network



Base Network



Standard linear regression models

	Idiosyncratic Network model			Morphology Network model			Base Network model		
	Estimate	*SE*		*Estimate*	*SE*		*Estimate*	*SE*	
Intercept	0.216901	0.026210	***	0.090708	0.025887	***	0.408246	0.029999	***
MEAN WORD SUPPORT	0.170726	0.023507	***	0.250262	0.020700	***	0.050723	0.012716	***
PATH ENTROPIES	-0.008688	0.002242	***	-0.008442	0.002309	***	-0.009342	0.002259	***
SEMANTIC DENSITY	-0.043545	0.008925	***	0.033868	0.012372	**	-0.093906	0.025844	***
SPEECH RATE	-0.058757	0.001148	***	-0.058602	0.001159	***	-0.058702	0.001171	***
N	4448			4456			4456		
R^2 *adjusted*	0.3778			0.3742			0.3623		

Mixed-effects regression models

	Idiosyncratic Network model			Morphology Network model			Base Network model		
	Estimate	*SE*		*Estimate*	*SE*		*Estimate*	*SE*	
Intercept	1.328e-01	4.601e-02	**	2.146e-01	6.024e-02	***	2.595e-01	2.510e-02	***
MEAN WORD SUPPORT	2.722e-01	4.600e-02	***	2.535e-01	4.572e-02	***	1.211e-01	2.654e-02	***
PATH ENTROPIES	-1.173e-02	5.625e-03	*	-1.163e-02	5.633e-03	*			
SEMANTIC VECTOR LENGTH	-1.606e-02	6.860e-03	*	-3.294e-02	1.550e-02	*			
SPEECH RATE	-5.944e-02	1.116e-03	***	-5.937e-02	1.116e-03	***	-5.936e-02	1.117e-03	***
N	4357			4358			4357		
R^2 *marginal*	0.3690016			0.3638608			0.3487138		

Traditional models

	Traditional standard regression model			Traditional mixed-effects model		
	Estimate	*SE*		*Estimate*	*SE*	
Intercept	3.888e-01	8.345e-03	***	4.159e-01	1.106e-02	***
WORD FREQUENCY	4.970e-08	3.764e-08		-2.608e-07	2.328e-07	
RELATIVE FREQUENCY	-2.136e-05	4.166e-05		-1.446e-05	8.931e-05	
BIGRAM FREQUENCY	-6.542e-07	6.293e-07		7.978e-07	6.382e-07	
MEAN BIPHONE PROBABILITY	-5.188e+00	8.872e-01	***	-7.167e+00	1.545e+00	***
AFFIX ation						
dis	8.145e-03	6.700e-03		-1.405e-03	1.438e-02	
ize	-2.316e-02	5.251e-03	***	-1.491e-02	1.377e-02	
less	-5.749e-02	8.226e-03	***	-7.569e-02	1.524e-02	***
ness	-5.473e-02	5.700e-03	***	-3.630e-02	1.295e-02	**
SPEECH RATE	-5.893e-02	1.163e-03	***	-5.986e-02	1.116e-03	***
N	4450			4354		
R^2 *adjusted/marginal*	0.3731			0.3705799		
R^2 *conditional*				0.5344904		

Relative importance of variables

	Relative importance metrics (lmg)							
	Idiosyncratic Network		Morphology Network		Base Network		Traditional model	
	lm	lmer	lm	lmer	lm	lmer	lm	lmer
MEAN WORD SUPPORT	0.0089	0.1649	0.0148	0.0956	0.0025	0.1641		
PATH ENTROPIES	0.0023	0.0031	0.0023	0.0017	0.0030			
SEMANTIC DENSITY	0.0067		0.0020		0.0014			
SEMANTIC VECTOR LENGTH		0.0064		0.0399				
SPEECH RATE	0.3605	0.1946	0.3556	0.2266	0.3559	0.1845	0.3561	0.2140
WORD FREQUENCY							0.0007	0.0065
RELATIVE FREQUENCY							0.0006	0.0044
BIGRAM FREQUENCY							0.0007	0.0034
MEAN BIPHONE PROBABILITY							0.0025	0.1178
AFFIX							0.0136	0.0246
total variance explained	0.3778	0.3690	0.3742	0.3639	0.3623	0.3487	0.3731	0.3706

Extract from closest semantic neighbors of DIS words

Word	Phones	Neighbors						
Idiosyncratic Network								
disarm	dIs,m	m1d1	kInt	w{m	m{mb5	kr{Nkl	n5zl	bl{NkIl
disband	dIsb{nd	m1d1	kInt	bl{NkIl	w{m	m{mb5	kr{Nkl	pIpIn
discard	dIsk,d	dIs@r1	dIst1st	dIskrEdIt	dIsgr1s	dIskVmf@t	$l	dIs@b1
discharge	dIsJ,=	dIsl2k	dIsQnIst	dIstrVst	dIs@gri	dIskVmf@t	dIsgr1s	dIsk@ntEnt
disclose	dIskl5z	m1d1	kInt	m{mb5	w{m	bl{NkIl	n5zl	SIt
discount	dIsk6nt	dIsQnIst	dIskVmf@t	dIsgr1s	dIsk@ntEnt	dIstrVst	dIst1st	dIsg2z
discourse	dIsk$s	dIs@r1	dIst1st	dIskrEdIt	dIsgr1s	dIskVmf@t	dIsp{r@tl	dIslQ=
disease	dIziz	dIskVv@R	dIs@p7R	dIs$d@R	dIsJ,=	dIsl2k	dIsk6nt	dIs@gri
disgrace	dIsgr1s	dIst1st	dIskVmf@t	dIs@r1	dIskrEdIt	dIs@b1	dIslQ=	dIsp{r@tl
Morphology Network								
disarm	dIs,m	dIsjun@tl	dIs5n	dIsb{nd	dIs@r1	dIskrEdIt	dIsp{r@tl	dIs@b1
disband	dIsb{nd	dIsjun@tl	dIs5n	dIs,m	dIs@r1	dIskrEdIt	dIs@b1	dIsp{r@tl
discard	dIsk,d	dIskVmf@t	dIsgr1s	dIst1st	dIsQnIst	dIs@r1	dIsk@ntEnt	dIslQ=
discharge	dIsJ,=	dIsl2k	dIsQnIst	dIstrVst	dIs@gri	dIskVmf@t	dIsgr1s	dIsk@ntEnt
disclose	dIskl5z	dIs@r1	dIs5n	dIs,m	dIskrEdIt	dIsjun@tl	dIsb{nd	dIsp{r@tl
discount	dIsk6nt	dIskVmf@t	dIsQnIst	dIsgr1s	dIsl2k	dIs@gri	dIstrVst	dIsg2z
discourse	dIsk$s	dIskVmf@t	dIsgr1s	dIst1st	dIsQnIst	dIsk@ntEnt	dIs@r1	dIsrIg,d
disease	dIziz	dIskVv@R	dIs@p7R	dIs$d@R	dIsJ,=	dIsl2k	dIsk6nt	dIs@gri
disgrace	dIsgr1s	dIst1st	dIskVmf@t	dIs@r1	dIskrEdIt	dIs@b1	dIslQ=	dIsp{r@tl
Base Network								
disarm	dIs,m	dIsg2z	dIsp{r@tl	dIsgVst	dIs@r1	dIsl2k	dIs@bidj@ns	dIspl1s
disband	dIsb{nd	dIsg2z	dIsp{r@tl	dIs@r1	dIsgVst	dIsl2k	dIspl1s	dIs@bidj@ns
discard	dIsk,d	dIsg2z	dIsp{r@tl	dIsgVst	dIs@r1	dIs@bidj@ns	dIsl2k	dIspl1s
discharge	dIsJ,=	dIsg2z	dIsp{r@tl	dIsgVst	dIs@r1	dIsl2k	dIs@bidj@ns	dIsQnIst
disclose	dIskl5z	dIsg2z	dIsp{r@tl	dIsgVst	dIs@r1	dIsl2k	dIs@bidj@ns	dIslQ=
discount	dIsk6nt	dIsg2z	dIsp{r@tl	dIsgVst	dIs@r1	dIs@bidj@ns	dIsl2k	dIsQnIst
discourse	dIsk$s	dIsg2z	dIsp{r@tl	dIsgVst	dIs@r1	dIspl1s	dIs@bidj@ns	dIsQnIst
disease	dIziz	dIsg2z	dIsp{r@tl	dIsgVst	dIs@r1	dIs@bidj@ns	dIsl2k	dIsQnIst
disgrace	dIsgr1s	dIsg2z	dIsp{r@tl	dIsgVst	dIs@r1	dIsl2k	dIs@bidj@ns	dIsQnIst

Problem

Many morpheme-based models of the morphology-phonology interaction and of speech production cannot account for morpho-phonetic findings.

- no post-lexical access to morphological information
- erasure of morphological brackets
- generic phoneme templates

e.g., Chomsky & Halle 1968, Kiparsky 1982, Dell 1986, Levelt et al. 1999, Roelofs & Ferreira 2019

Future directions

We think it could be worthwhile to…

- analyze durations for a larger dataset with more derivational functions.
- train lexome-to-lexome vectors without coding for function lexomes in the first place.
- explore how to build vectors for words with multiple derivational functions.