Individual Differences and Stress Variation in English Complex Words

Tammy Ganster, Trier University (Germany)

ganster@uni-trier.de

56th Annual Meeting of the *Societas Linguistica Europaea*

National and Kapodistrian University of Athens

Aug 29 - Sep 1 2023

UNIVERSITÄT TRIER

idéntifiable	identifíable
tríumphant	triúmphant
eváluative	evaluátive
discríminatory	discriminátory

Variable stress in English complex adjectives

embedded base	stress preserving derivative variant	stress non-preserving derivative variant
i**dén**tify	i**dén**tifiable	identi**fí**able
tríumph	**trí**umphant	tri**úm**phant
e**vál**uate	e**vál**uative	evalu**á**tive
dis**crí**minate	dis**crí**minatory	discrimi**ná**tory

The phenomenon: primary stress of embedded base is not always preserved within the same morphological category

What factors can account for this variation?

Previous research on stress variability

Stratum-based approaches (Kiparsky 1982 et seq., 2005, 2015, Fudge 1984)	**Structural** approaches (Trevian 2003, 2007, Bauer, Lieber & Plag 2013, Newell 2020)	**Paradigmatic** approaches (Steriade 1999a, 1999b, Stanton & Steriade 2014, 2021, Breiss 2021)
▶ strict division into three categories ▶ **stress shifting (stem level): -**ory > óscillate > oscillátory, oscíllatory, *óscillatory ▶ **stress preserving (word level):** -ness > alért > alértness, *álertness ▶ **variable (dual level):** -able > jústify > jústifiable ~ justifíable	▶ **segmental phonological features** of derivative assumed to influence stress position ▶ célebrate > célebr[ə]tory ~ celebr[éɪ]tory	▶ suggests **broader paradigmatic relationships** may be at play ▶ **embedded base** is **not considered the only influential base** ▶ ~~démonstrate~~ demónstrative > demónstrable ?

Issues of structure-based accounts

Stratum-based accounts

- empirical **evidence against uniform behavior of stem and word level**

→ stress preservation as well as stress shift underpredicted (*oscillátory ~ óscillatory*)

Structural accounts

- **effect** said to be **variable** (*ánalyze > ánal[aɪ]zable ~ anal[áɪ]zable*)
- can indentify a possible reason for stress shift but **cannot account for variability of stress shift**

Paradigmatic approaches

- does not take full extent of **morphological relationships** in paradigm into account
- **does not take psycholinguistic factors** such as lexical frequency seriously enough

- **all accounts purely phonological even though we are talking about a morpho-phonological phenomenon**

Processing-based accounts: a possible solution?

- based on **Hay's dual-route model of lexical access** (Hay 2001, 2003, Hay & Baayen 2003)
- relationship between **derivative frequency and base frequency** is crucial

- if derivative frequency **< embedded base frequency** → **decomposition route**
  - *anticipatory* = *antícipate* + *-ory* → *antícipatory*
  - **stress preserved**
- if **derviative frequency >** embedded base frequency → **whole-word route**
  - *derogatory* = *derogatory* ~~*(dérogate + -ory)*~~ → *derogátory*
  - **stress not preserved**

- see Collie 2007, 2008, Bermúdez-Otero 2012, Dabouis 2017 for pertinent studies based on this approach

Issues of processing-based accounts

- processing-based accounts have also not fully implemented all assumptions that come with taking a morphological approach such as **individual differences in processing**
  - exclusive **reliance on corpus frequencies to account for processing effects**
  - **not compatible with individual differences**
  - previous studies (Arndt-Lappe & Sanz 2017; Ganster 2019) have shown that not all speakers reflect corpus frequency in the same way
- **individual differences generally marginalized**, to date mainly only studied in reading acquistion (e.g. McCutchen et al. 2009) and second language acquisition (e.g. Coxhead et al. 2015)

Towards more individual models of morphological processing

- **more individual measures of morpholgical processing are needed**
- correlates that emerged as important in other fields
  - individual **awareness of morphological structure** (e.g. McCutchen et al. 2009)
  - **vocabulary size** (e.g. Brysbaert et al. 2016, Mainz et al. 2017)
- These measures **need to be tested for their effects on stress production in complex words**

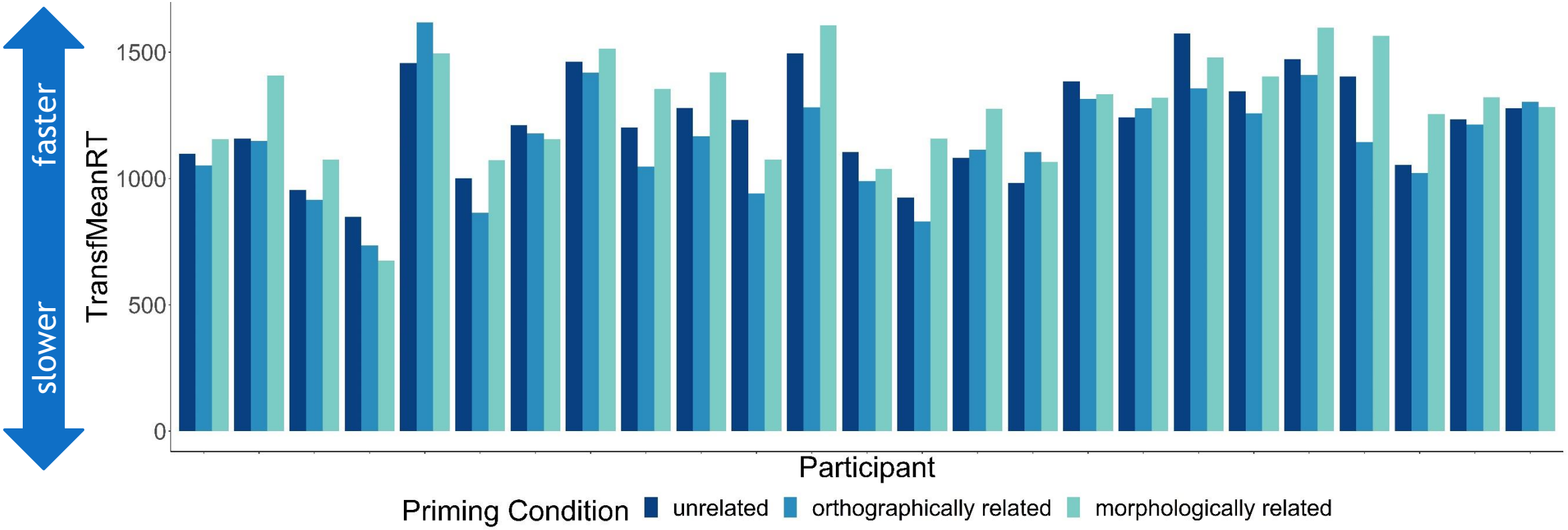
Experiment

- **remote online** experiment
- **153 native speakers** of British English
  - age: **18-77 yrs,** mean: 25, median: 29.98 / **93 females, 60 males**
  - **recruited via the online platform Prolific**
- **multi-task experiment**
  - **PROCESSING**
    - **morphological sensitivity** task (masked priming with lexical decision)
    - **vocabulary size** test (standardized test, Coxhead et al. 2015, Nation & Beglar 2007)
  - **SOCIO-DEMOGRAPHIC DATA**
    - meta questionnaire (education, languages, geography, socio-economic status...)
  - **STRESS**
    - **production task (read out test sentences with complex adjectives in them)**
    - perception task (imitation task)

Data overview

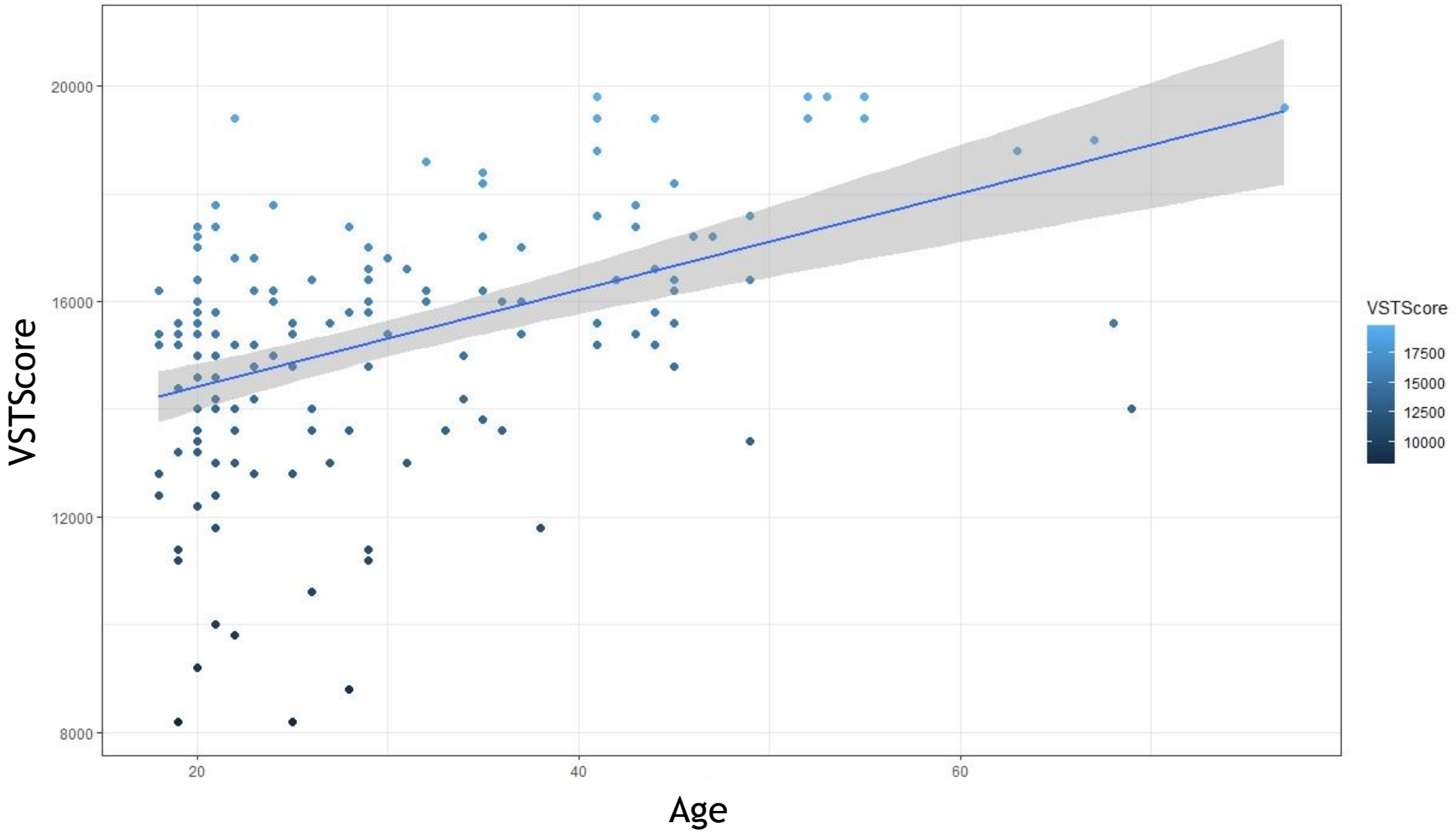
morphological sensitivity task	vocabulary size test	production task
3,467 observations	15,300 observations	
masked priming experiment with **lexical decision task**	**standardized** and multiply validated **forced choice test** (Nation & Beglar 2007)	
measured **reaction time** to **three different priming conditions**:	**100 questions**	
complex words **primes - simplex** words **targets**	**Example**	
▶ **morphologically related** prime-target pairs *(subversion - subvert)*	*see: They SAW it.* *a) closed it tightly* *b) waited for it* *c) looked at it*	will be explained later
▶ **orthographically related** prime-target pairs *(chargeable - charisma)*		
▶ **unrelated** prime-target pairs *(inventive - remorse)*	**score from 0 - 20,000** (**estimates number of known word families**)	

Individuality in morphological sensitivity

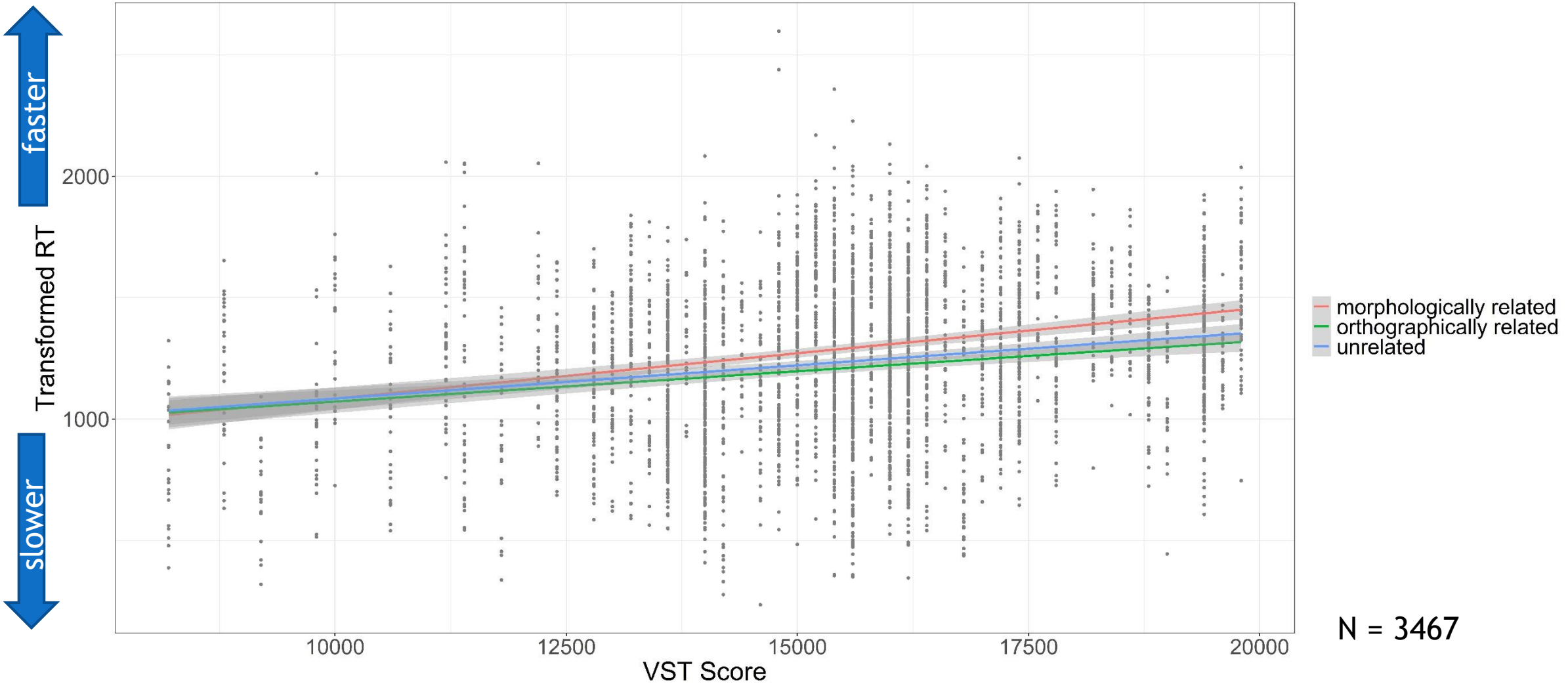


random sample of 25 out of 129 participants
Separate model: lmer(RTtr ~ Condition + VSTScore + (1+ Condition | Participant) , data = MSfinal)
→ showed priming condition has significant effect on RT

Variation in vocabulary size



Variation in the effect of vocabulary size



N = 3467

Relating individual differences in morphological processing to stress variation

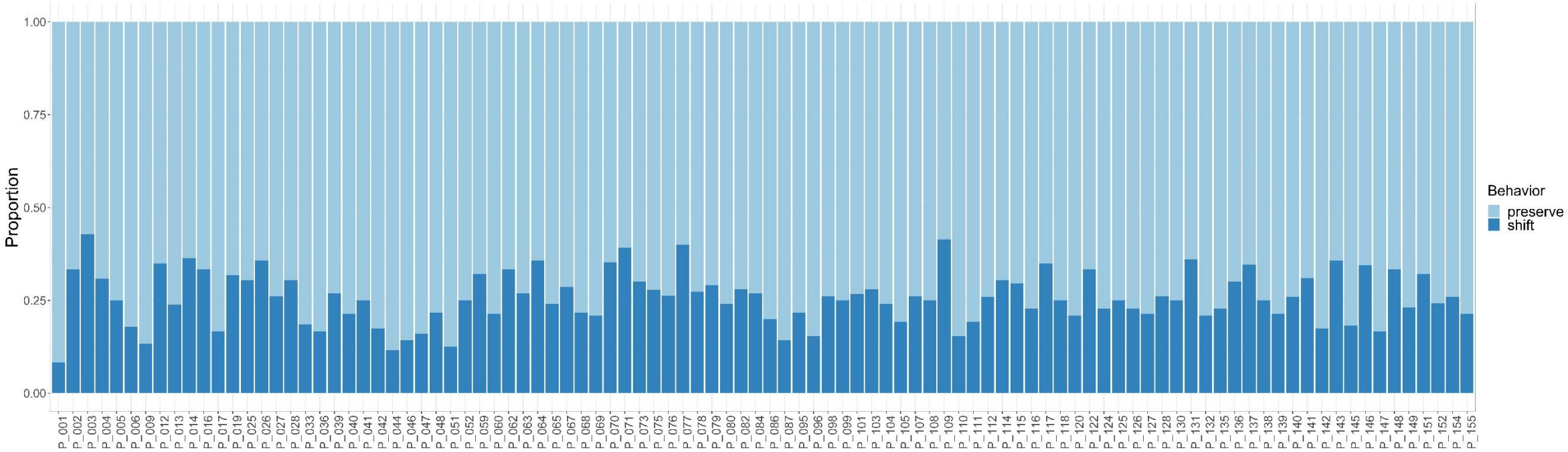
- **Hypothesis 1: greater sensitivity to morphological structure**
  - more decomposition

    → **more stress preservation**
- **Hypothesis 2: greater vocabulary size**
  - more decomposition → **more preservation?**
  - paradigmatic effects → **paradigmatically dominant stress preferred?**

Experiment: Stress Production

- **production task**
- **3,400** observations from 153 native speakers of British English
- test sentences from Corpus of American Soap Operas (Davies 2011) with complex ***-able, -ant , -ive, -ory*** **adjectives**
- each participant read out **30 test sentences**
  - *We're trying to do something a little more innovative.*
  - *Well, actually, this move was anticipatory.*
  - *Is there an address or a name or something else that's identifiable?*
- each recording **assessed by three trained raters** (raters **agree in 77% of cases, only agreement cases taken into account** in analyses)

Individuality in stress variation



For all 98 participants who have lived in the UK all their lives

N = 2316

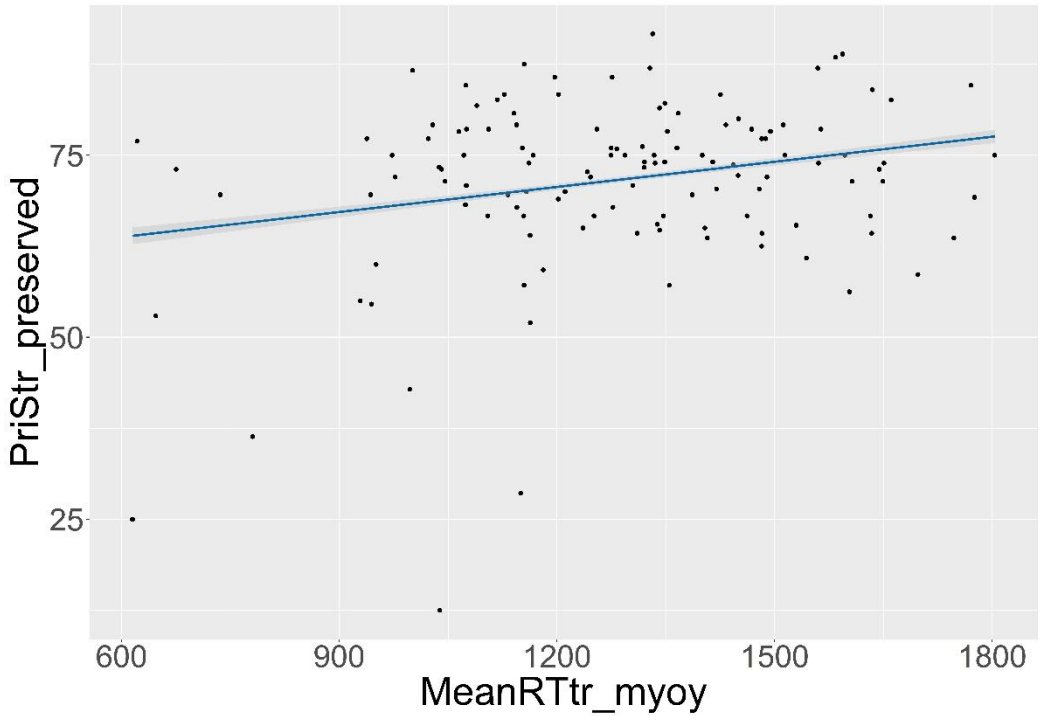
min proportion of non-preservation = **8%**

max proportion of non-preservation = **43%**

SD in proportion of non-preservation = **7%**

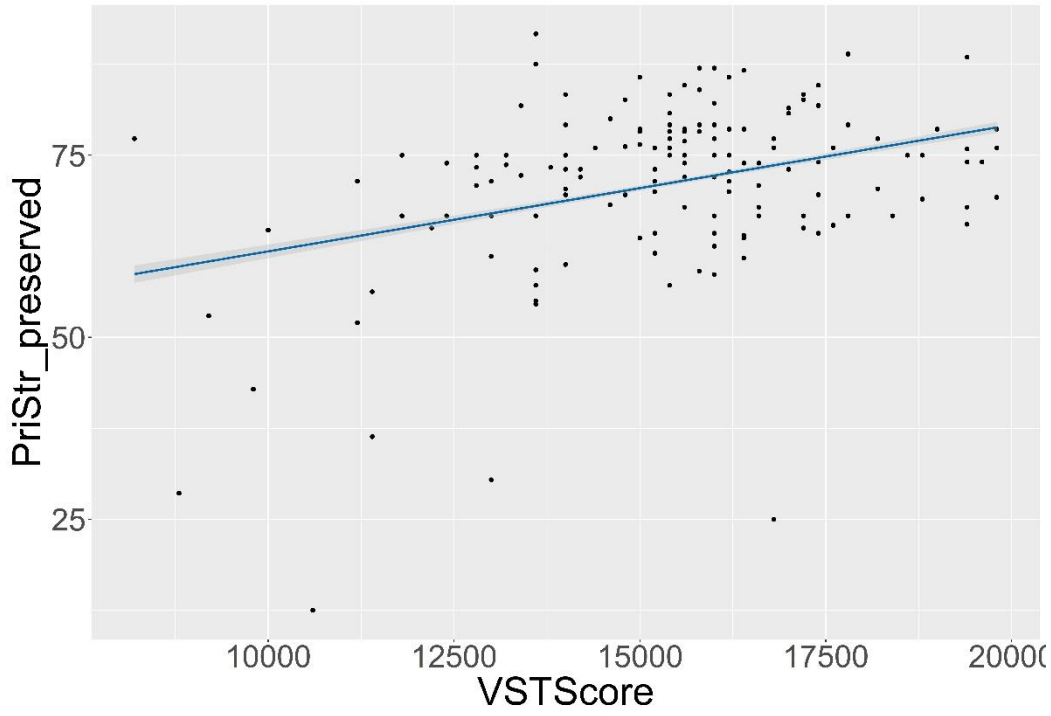
Individual processing and primary stress preservation

N = 2919



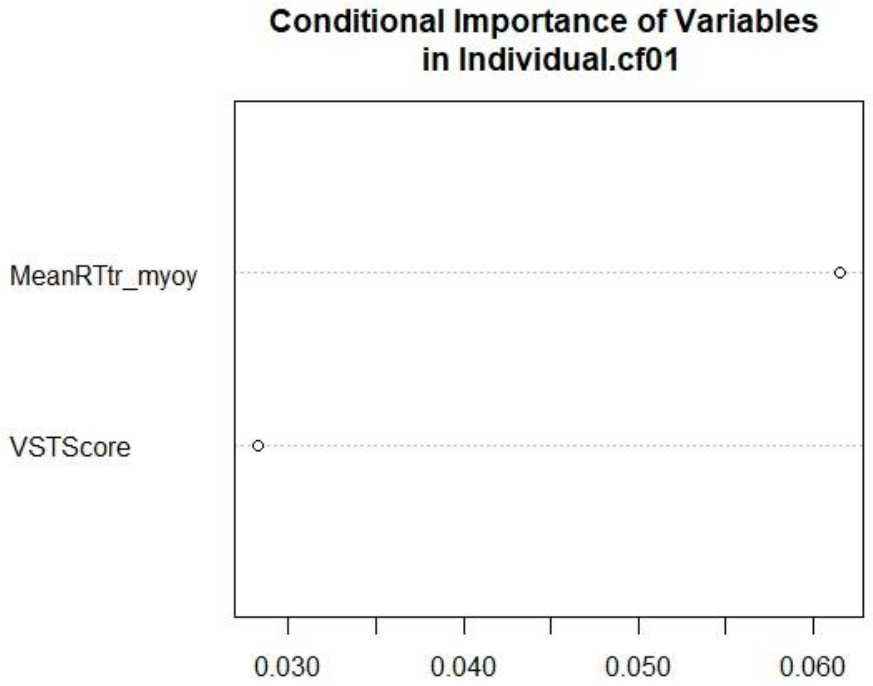
slower faster

N = 3442



Individual processing and primary stress preservation

- transformed **mean reaction time** to morphologically related prime-target pairs and **vocabulary size test score** are **moderately correlated**
  - correlation coefficient: 0.34
  - condition number: 19 (moderate correlation)
- **conditional random forest analysis**
  - less sensitive to correlated predictors
  - cforest::partykit (Hothorn & Zeileis 2015; Hothorn et al. 2006, Zeileis et al. 2008)
- OOB accuracy of cforest 0.72

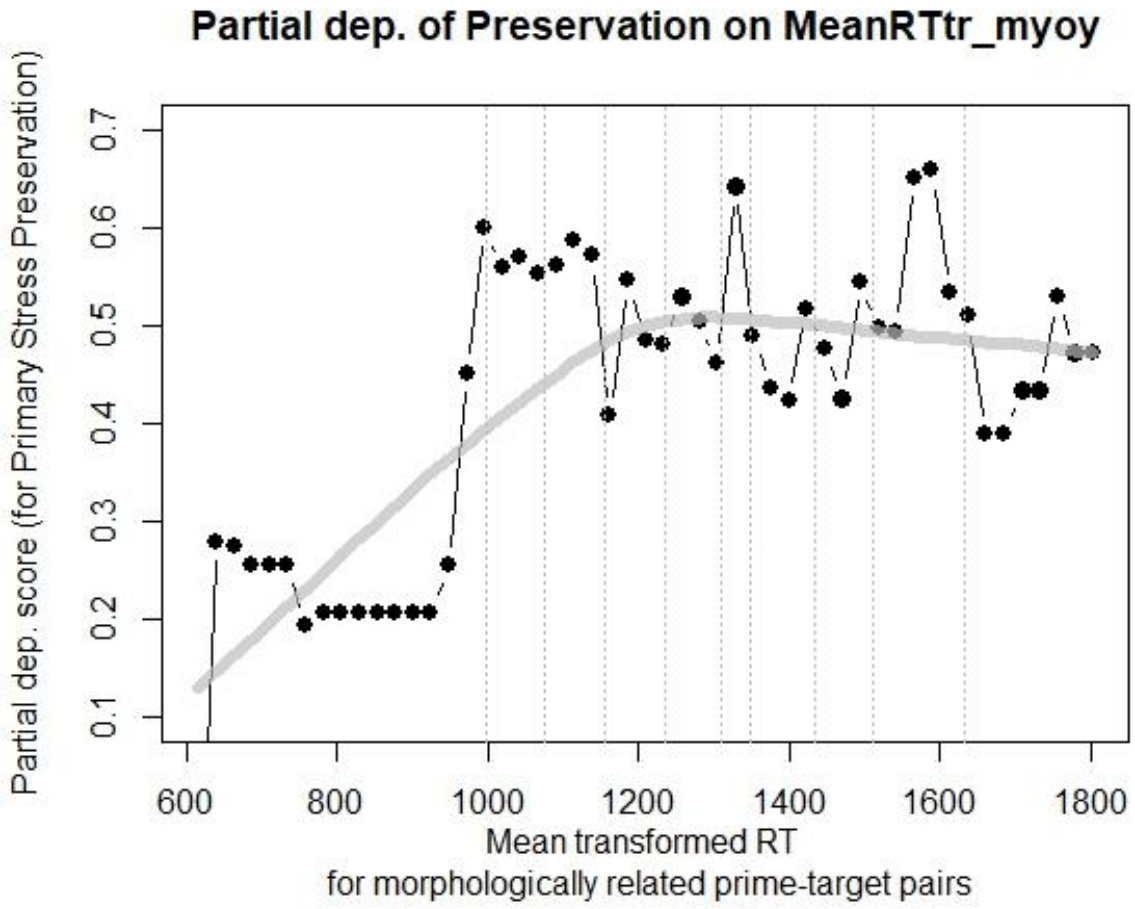


PriStrPreservation ~ MeanRTtr_myoy + VSTScore, data = AS_RTs, ntree = 500, perturb=list(replace=TRUE), mtry = 2

non-conditional variable importance yielded same result

Morphological Sensitivity and Stress Preservation

- partial dependence scores can be used to determine the **nature and direction of effects** in a random forest model
- see Gries 2021: Chpt 7 for more information



Conclusion

- **individual differences** in morphological processing, vocabulary size and stress placement preferences
- **larger vocabulary size facilitates** morphological processing
- **morphological sensitivity** emerged as **more important predictor**
  - partial dependence scores of random forest model indicate **a faster reaction time to morphologically related prime-target pairs boosts stress preservation in speakers**
  - **more exploration and validation with other models needed**

Thank you/Ευχαριστώ πολύ
for your attention!

Feel free to contact me: ganster@uni-trier.de

References

Arndt-Lappe, S., & Sanz, J. (2017, June 21). *Stress Variability in English -able and -ory Adjedctives: Markedness, Faithfulness, and Usage*. 11th Mediterranean Morphology Meeting, Nicosia, Cyprus.

Baayen, H., & Milin, P. (2010). Analyzing Reaction Times. *International Journal of Psychological Research*, *3*(2), 12–28.

Bauer, L., Lieber, R., & Plag, I. (2013). *The Oxford Reference Guide to English Morphology*. Oxford University Press.

Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). How Many Words Do We Know? Practical Estimates of Vocabulary Size Dependent on Word Definition, the Degree of Language Input and the Participant's Age. *Frontiers in Psychology*, *7*.

Collie, S. (2008). English stress preservation: The case for "fake cyclicity." *English Language and Linguistics*, *12*(3), 505–532.

Coxhead, A., Nation, P., & Sim, D. (2015). Measuring the Vocabulary Size of Native Speakers of English in New Zealand Secondary Schools. *NZ J Educ Stud*, *50*, 121–135.

Davies, M. (2011). *Corpus of American Soap Operas*. https://www.english-corpora.org/scotus/

Fudge, E. (1984). *English Word Stress*. G. Allen & Unwin.

Gries, S. Th. (2021). *Statistics for Linguistics With R* (3rd ed.). De Gruyter Mouton.

Hay, J. (2001). Lexical frequency in morphology: Is everything relative? *Linguistics*, *39*(6), 1041–1070.

Hay, J. (2003). *Causes and Consequences of Word Structure*. Routledge.

Hothorn, T., Hornik, K., & Zeileis, A. (2006). Unbiased Recursive Partitioning: A Conditional Inference Framework. Journal of Computational and Graphical Statistics, 15(3), 651–674. https://doi.org/doi:10.1198/106186006X133933

Hothorn, T., & Zeileis, A. (2015). partykit: A Modular Toolkit for Recursive Partytioning in R. Journal of Machine Learning Research, 16, 3905–3909.

Kiparsky, P. (2005). *Paradigm Uniformity Constraints*. https://web.stanford.edu/~kiparsky/Papers/LexConservatism.pdf

Kiparsky, P. (2015). Stratal OT: A synopsis and FAQs. In Y. E. Hsiao & L. H. Wee (Eds.), *Capturing Phonological Shades* (pp. 2–44). Cambridge Scholars Publishing.

References

Mainz, N., Shao, Z., Brysbaert, M., & Meyer, A. S. (2017). Vocabulary Knowledge Predicts Lexical Processing: Evidence from a Group of Participants with Diverse Educational Backgrounds. *Frontiers in Psychology*, *8*(1164).

McCutchen, D., Logan, B., & Biangardi-Orpe, U. (2009). Making Meaning: Children's Sensitivity to Morphological Information During Word Reading. *Reading Research Quarterly*, *44*(4), 360–376.

Nation, P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, *31*(7).

Stanton, J., & Steriade, D. (2021). *Markedness drives base selection: Experimental evidence*. The 28th Manchester Phonology Meeting, Manchester, England.

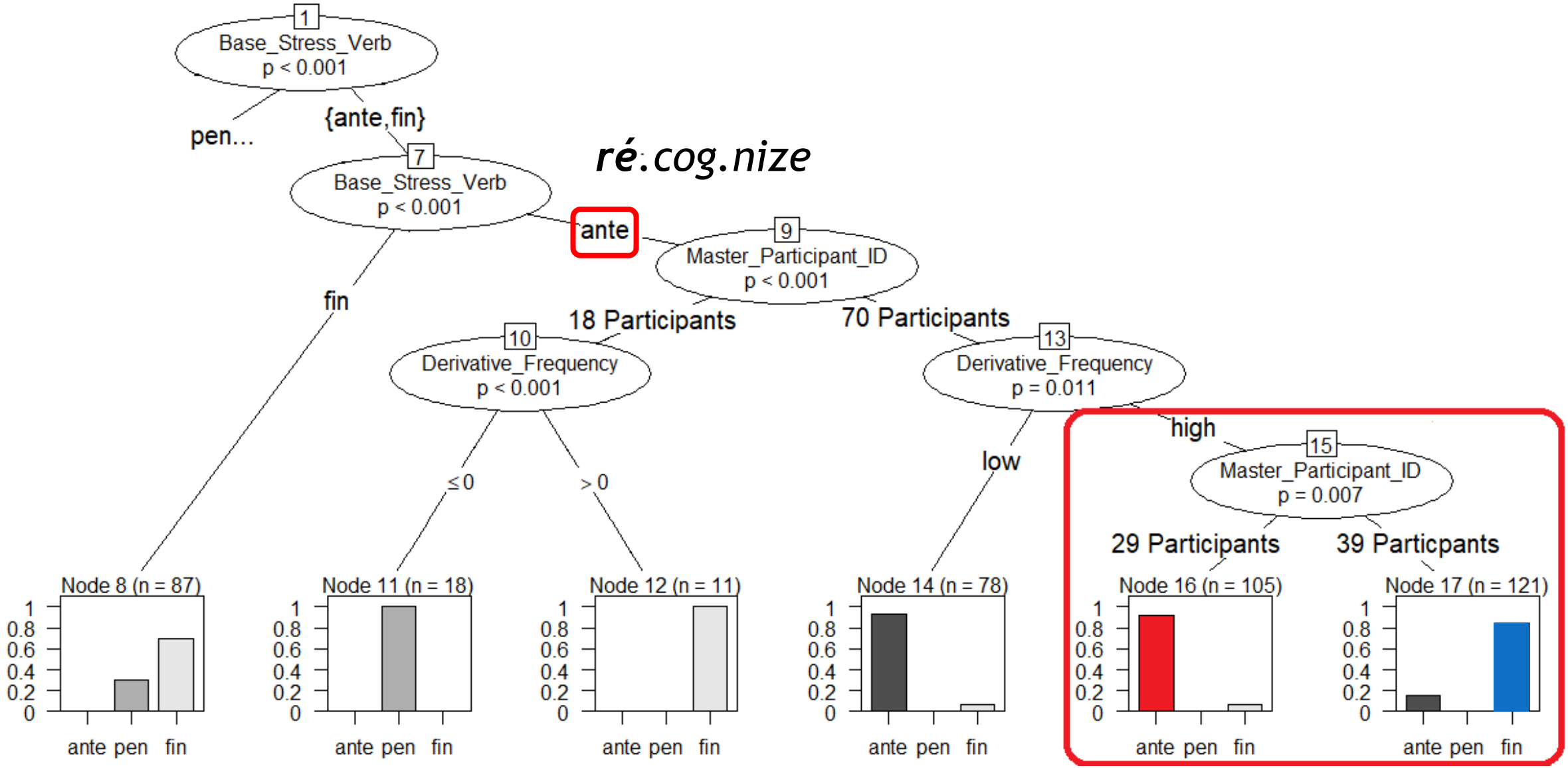
Steriade, D. (1999a). Lexical Conservatism. *Linguistics in the Morning Calm - Selected Papers from SICOL 1997*, 157–179.

Steriade, D. (1999b). Lexical Conservatism in French Adjectival Liaison. In B. Bullock, M. Authier, & L. Reed (Eds.), *Formal Perspectives in Romance Linguistics* (pp. 243–270). John Benjamins Publishing Company.

Trevian, I. (2003). *Morphoaccentologie et processus d'affixation de l'anglais*. Peter Lang.

Trevian, I. (2007). Stress-neutral endings in contemporary British English: An updated overview. *Language Sciences*, *29*, 426–450.

Zeileis, A., Hothorn, T., & Hornik, K. (2008). Model-Based Recursive Partitioning. Journal of Computational and Graphical Statistics, 17(2), 492–514. https://doi.org/doi:10.1198/106186008X319331



***ré*.cog.nize**

***ré*.cog.niz.able ~ re.cog.*níz*.able**

(N = 711, 31 types; goodness-of-fit = 90%, from Ganster 2019, see also Arndt-Lappe & Sanz 2017 for complementary effect with base frequency)