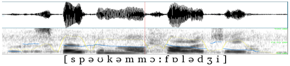
Co-articulation between stem vowels and suffixes: semantics all the way down

Motoki Saito, Fabian Tomaschek, R. Harald Baayen

Words in the World

18.10.2020

Heinrich Heine Universität Düsseldorf



EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Frequency effect in speech production

- Lemma frequency?? (Levelt et al., 1999)

Frequency effect in speech production

- Lemma frequency?? (Levelt et al., 1999)

  v.s.

- Word frequency?? (Janssen et al., 2008)

Frequency effect in speech production

- Lemma frequency?? (Levelt et al., 1999)

  v.s.

- Word frequency?? (Janssen et al., 2008)

$\downarrow$

- Contrasting evidence

Frequency effect in speech production

- Lemma frequency?? (Levelt et al., 1999)

  v.s.

- Word frequency?? (Janssen et al., 2008)

$\downarrow$

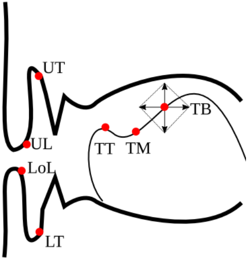
- Contrasting evidence

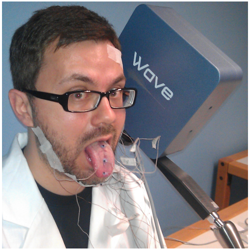
$\Downarrow$

- Tongue movements
  - $\Rightarrow$ New opportunities to study cognitive process driving speech process.

Electromagnetic Articulography (EMA)

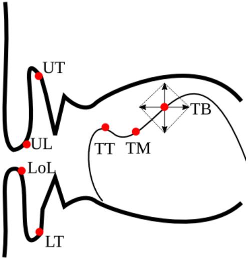
- Sensors glued on the tongue.

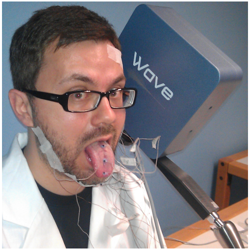




Electromagnetic Articulography (EMA)

- Sensors glued on the tongue.
- Sensor positions are registered separately.





Previous result (Tomaschek, Tucker, et al., 2018)

Previous result (Tomaschek, Tucker, et al., 2018)

- Ariculations of a stem vowel [aː], e.g. *sagt*, depend on...

Previous result (Tomaschek, Tucker, et al., 2018)

- Ariculations of a stem vowel [aː], e.g. *sagt*, depend on...
  - not only **suffixes**, e.g. *sagt*

Previous result (Tomaschek, Tucker, et al., 2018)

- Ariculations of a stem vowel [aː], e.g. *sagt*, depend on...
  - not only **suffixes**, e.g. *sagt*
  - but also **frequency**

Previous result (Tomaschek, Tucker, et al., 2018)

- Ariculations of a stem vowel [aː], e.g. *sagt*, depend on...
  - not only **suffixes**, e.g. *sagt*
  - but also **frequency**

    $\rightarrow$ **Practice** effect on articulation

Word frequency or Syllable frequency

- The practice effect is driven by...
  - word frequency?? (Janssen et al., 2008)
  - syllable frequency?? (Levelt et al., 1999)

Experimental design

- Keep syllables constant

Experimental design

- Keep syllables constant
- Vary morphological conditions

Experimental design

- Keep syllables constant
- Vary morphological conditions

  e.g. *geschafft* [gəʃaft] “made/managed it”.

Experimental design

- Keep syllables constant
- Vary morphological conditions

  e.g. *geschafft* [gəʃaft] “made/managed it”.
  v.s.
  Fachschaft [faxʃaft] “student association of a (university) department”.

Experimental design

- Keep syllables constant
- Vary morphological conditions

e.g. *geschafft* [gə**ʃaft**] "made/managed it".
v.s.
Fachschaft [fax**ʃaft**] "student association of a (university) department".

Experimental design

- Keep syllables constant
- Vary morphological conditions

e.g. *ge-schaff-t* [gə-**ʃaf-t**] "made/managed it".
v.s.
Fach-schaft [fax-**ʃaft**] "student association of a (university) department".

Methodological question

- How can we best predict what the tongue is doing??

Methodological question

- How can we best predict what the tongue is doing??
- Classical view:
  - words, lemmas, morphemes, syllables, phones, ...

Methodological question

- How can we best predict what the tongue is doing??
- Classical view:
  - words, lemmas, morphemes, syllables, phones, ...
- How about **semantics**??

Methodological question

- How can we best predict what the tongue is doing??
- Classical view:
  - words, lemmas, morphemes, syllables, phones, ...
- How about **semantics**??

Methodological question

- How can we best predict what the tongue is doing??
- Classical view:
  - words, lemmas, morphemes, syllables, phones, ...
- How about **semantics**??

$\Downarrow$

- Alternative:

Methodological question

- How can we best predict what the tongue is doing??
- Classical view:
  - words, lemmas, morphemes, syllables, phones, ...
- How about **semantics**??

$\Downarrow$

- Alternative:
  - Production driven by semantics

Methodological question

- How can we best predict what the tongue is doing??
- Classical view:
  - words, lemmas, morphemes, syllables, phones, ...
- How about **semantics**??

$\Downarrow$

- Alternative:
  - Production driven by semantics
  - Computational model that predicts forms from semantics.

Linear Discriminative Learning (LDL) models (R. H. Baayen et al., 2019)*

- Simple 2-layer network which...

*LDL available in R (WpmWithLdl)

Linear Discriminative Learning (LDL) models (R. H. Baayen et al., 2019)*

- Simple 2-layer network which...
  - can map forms onto meanings

*LDL available in R (WpmWithLdl)

Linear Discriminative Learning (LDL) models (R. H. Baayen et al., 2019)*

- Simple 2-layer network which...
  - can map forms onto meanings
  - can map meanings onto forms

*LDL available in R (WpmWithLdl)

Linear Discriminative Learning (LDL) models (R. H. Baayen et al., 2019)*

- Simple 2-layer network which...
  - can map forms onto meanings
  - can map meanings onto forms
  - is based on discriminative learning mechanism (Rescorla-Wagner learning rule)

*LDL available in R (WpmWithLdl)

Previous studies with LDL

- Duration of word final "S" (e.g. *plays*)
  - (Tomaschek, Plag, et al., 2019)

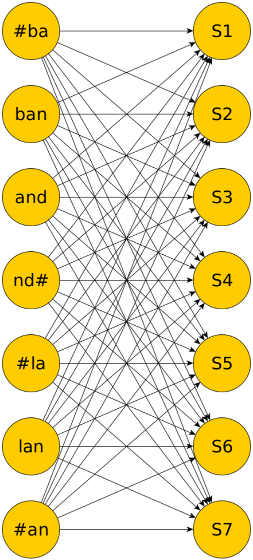
Previous studies with LDL

- Duration of word final "S" (e.g. *plays*)
  - (Tomaschek, Plag, et al., 2019)
- Duration of non-words
  - (Chuang et al., 2020)

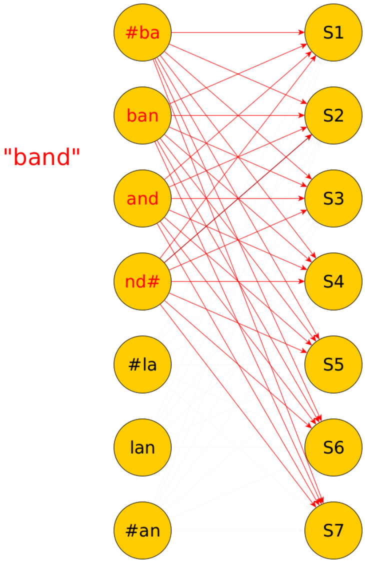
Previous studies with LDL

- Duration of word final "S" (e.g. *plays*)
  - (Tomaschek, Plag, et al., 2019)
- Duration of non-words
  - (Chuang et al., 2020)
- Segment duration at a morpheme boundary
  - (R. H. Baayen et al., 2019)

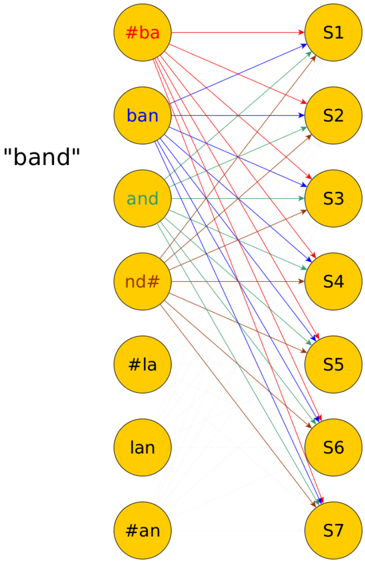
Functional load of triphones (1/2)



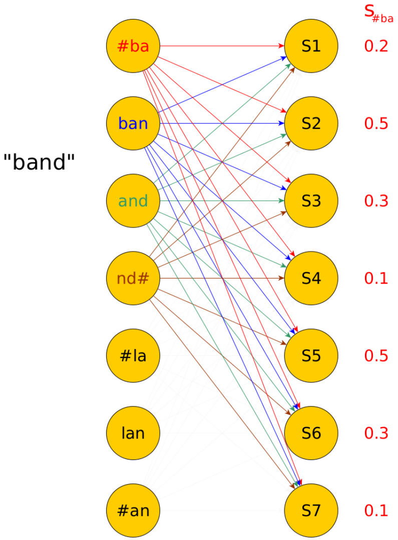
Functional load of triphones (1/2)



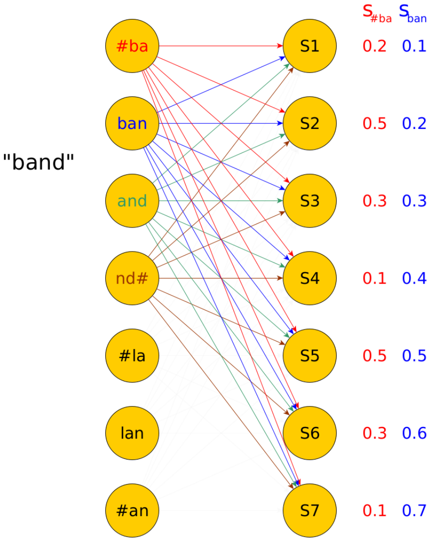
Functional load of triphones (1/2)



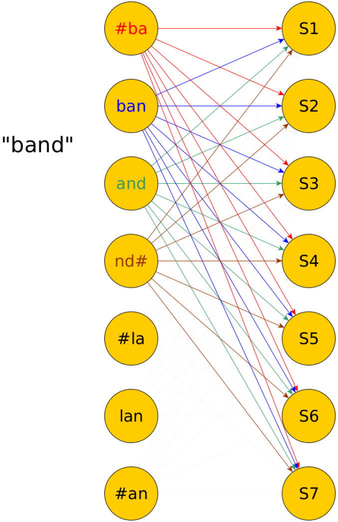
Functional load of triphones (1/2)



Functional load of triphones (1/2)

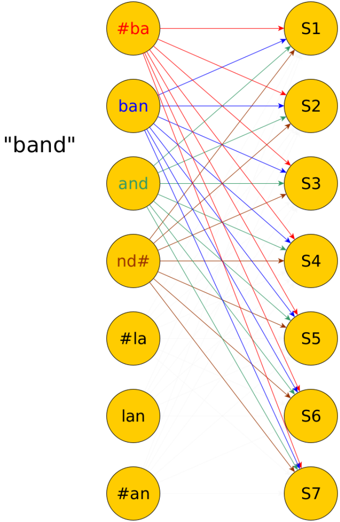


Functional load of triphones (1/2)



	$S_{\#ba}$	S_{ban}	S_{and}
S1	0.2	0.1	0.3
S2	0.5	0.2	0.1
S3	0.3	0.3	0.5
S4	0.1	0.4	0.7
S5	0.5	0.5	0.2
S6	0.3	0.6	0.3
S7	0.1	0.7	0.9

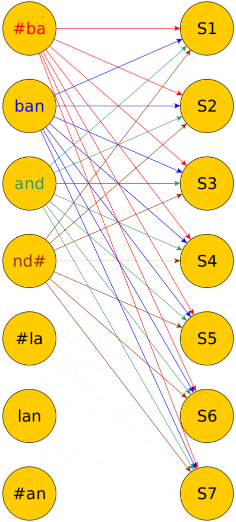
Functional load of triphones (1/2)



	$S_{\#ba}$	S_{ban}	S_{and}	$S_{nd\#}$
S1	0.2	0.1	0.3	0.8
S2	0.5	0.2	0.1	0.6
S3	0.3	0.3	0.5	0.1
S4	0.1	0.4	0.7	0.2
S5	0.5	0.5	0.2	0.4
S6	0.3	0.6	0.3	0.5
S7	0.1	0.7	0.9	0.2

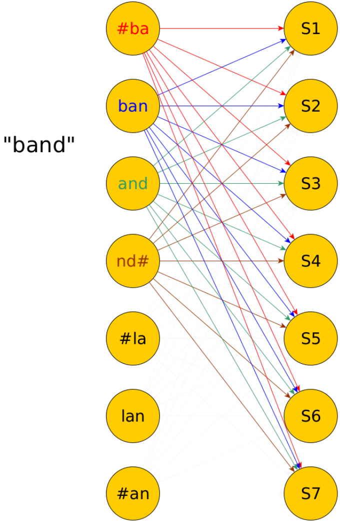
Functional load of triphones (1/2)

"band"



$S_{\#ba}$	S_{ban}	S_{and}	$S_{nd\#}$		$\hat{S}_{band}$
0.2	0.1	0.3	0.8	→	1.4
0.5	0.2	0.1	0.6	→	1.4
0.3	0.3	0.5	0.1	→	1.2
0.1	0.4	0.7	0.2	→	1.4
0.5	0.5	0.2	0.4	→	1.6
0.3	0.6	0.3	0.5	→	1.7
0.1	0.7	0.9	0.2	→	1.9

Functional load of triphones (1/2)



$S_{\#ba}$	S_{ban}	S_{and}	$S_{nd\#}$		$\hat{S}_{band}$	S_{band}
0.2	0.1	0.3	0.8	→	1.4	1.3
0.5	0.2	0.1	0.6	→	1.4	1.6
0.3	0.3	0.5	0.1	→	1.2	1.2
0.1	0.4	0.7	0.2	→	1.4	1.3
0.5	0.5	0.2	0.4	→	1.6	1.6
0.3	0.6	0.3	0.5	→	1.7	1.9
0.1	0.7	0.9	0.2	→	1.9	1.9

Functional load of triphones (2/2)

- $L_{\#ba} = corr(\mathbf{s}_{\#ba}, \mathbf{s}_{band})$

Functional load of triphones (2/2)

- $L_{\#ba} = corr(\mathbf{s}_{\#ba}, \mathbf{s}_{band})$
  - This measure...

Functional load of triphones (2/2)

- $L_{\#ba} = corr(\mathbf{s}_{\#ba}, \mathbf{s}_{band})$
  - This measure...
    - quantifies contribution of a specific triphone to the target semantic vector.

Functional load of triphones (2/2)

- $L_{\#ba} = corr(\mathbf{s}_{\#ba}, \mathbf{s}_{band})$
  - This measure...
    - quantifies contribution of a specific triphone to the target semantic vector.
    - tells us to what extent a particular triphone contributes to create the target semantic vector.

Functional load of triphones (2/2)

- $L_{\#ba} = corr(\mathbf{s}_{\#ba}, \mathbf{s}_{band})$
  - This measure...
    - quantifies contribution of a specific triphone to the target semantic vector.
    - tells us to what extent a particular triphone contributes to create the target semantic vector.
  - The higher L

Functional load of triphones (2/2)

- $L_{\#ba} = corr(\mathbf{s}_{\#ba}, \mathbf{s}_{band})$
  - This measure...
    - quantifies contribution of a specific triphone to the target semantic vector.
    - tells us to what extent a particular triphone contributes to create the target semantic vector.
  - The higher L

    $\rightarrow$ The more the target semantic vector is determined by this certain triphone.

Functional load of triphones (2/2)

- $L_{\#ba} = corr(\mathbf{s}_{\#ba}, \mathbf{s}_{band})$
  - This measure...
    - quantifies contribution of a specific triphone to the target semantic vector.
    - tells us to what extent a particular triphone contributes to create the target semantic vector.
  - The higher L
    - $\rightarrow$ The more the target semantic vector is determined by this certain triphone.

$\Downarrow$

- **Functional Load** of triphones

Relative functional load of triphones

- Stem triphones, e.g. *bemalt* [bəmaːlt].

Relative functional load of triphones

- Stem triphones, e.g. *bemalt* [bəmaːlt].
- Suffix triphones, e.g. *bemalt* [bəmaːlt].

Relative functional load of triphones

- Stem triphones, e.g. *bemalt* [bəmaːlt].
- Suffix triphones, e.g. *bemalt* [bəmaːlt].

$\Downarrow$

- Relative functional load: $L_{stem} - L_{suffix}$

Relative functional load of triphones

- Stem triphones, e.g. *bemalt* [bəmaːlt].
- Suffix triphones, e.g. *bemalt* [bəmaːlt].

$\Downarrow$

- Relative functional load: $L_{stem} - L_{suffix}$
  - Greater relative functional load

Relative functional load of triphones

- Stem triphones, e.g. *bemalt* [bəmaːlt].
- Suffix triphones, e.g. *bemalt* [bəmaːlt].

$\Downarrow$

- Relative functional load: $L_{stem} - L_{suffix}$
  - Greater relative functional load
    - $\rightarrow$ Stem is more important to get to the target meaning.

Variables: Classical Model

- Dependent variable:

Variables: Classical Model

- Dependent variable:
  - Tongue sensor positions

Variables: Classical Model

- Dependent variable:
  - Tongue sensor positions
- Main predictors:

Variables: Classical Model

- Dependent variable:
  - Tongue sensor positions
- Main predictors:
  - Time (normalized)

Variables: Classical Model

- Dependent variable:
  - Tongue sensor positions
- Main predictors:
  - Time (normalized)
  - Frequency (log)

Variables: Classical Model

- Dependent variable:
  - Tongue sensor positions
- Main predictors:
  - Time (normalized)
  - Frequency (log)
  - Morphological conditions

Variables: Classical Model

- Dependent variable:
  - Tongue sensor positions
- Main predictors:
  - Time (normalized)
  - Frequency (log)
  - Morphological conditions
  - Tongue types (e.g. tongue tip/body, vertical/horizontal)

Variables: Classical Model

- Dependent variable:
  - Tongue sensor positions
- Main predictors:
  - Time (normalized)
  - Frequency (log)
  - Morphological conditions
  - Tongue types (e.g. tongue tip/body, vertical/horizontal)
- Covariates:

Variables: Classical Model

- Dependent variable:
  - Tongue sensor positions
- Main predictors:
  - Time (normalized)
  - Frequency (log)
  - Morphological conditions
  - Tongue types (e.g. tongue tip/body, vertical/horizontal)
- Covariates:
  - Segment duration

Variables: Classical Model

- Dependent variable:
  - Tongue sensor positions
- Main predictors:
  - Time (normalized)
  - Frequency (log)
  - Morphological conditions
  - Tongue types (e.g. tongue tip/body, vertical/horizontal)
- Covariates:
  - Segment duration
- Random effects:

Variables: Classical Model

- Dependent variable:
  - Tongue sensor positions
- Main predictors:
  - Time (normalized)
  - Frequency (log)
  - Morphological conditions
  - Tongue types (e.g. tongue tip/body, vertical/horizontal)
- Covariates:
  - Segment duration
- Random effects:
  - Speaker

Variables: Classical Model

- Dependent variable:
  - Tongue sensor positions
- Main predictors:
  - Time (normalized)
  - Frequency (log)
  - Morphological conditions
  - Tongue types (e.g. tongue tip/body, vertical/horizontal)
- Covariates:
  - Segment duration
- Random effects:
  - Speaker
  - Previous segment, e.g. *sie sagt das* [ziː zaːkt das]

Variables: Classical Model

- Dependent variable:
  - Tongue sensor positions
- Main predictors:
  - Time (normalized)
  - Frequency (log)
  - Morphological conditions
  - Tongue types (e.g. tongue tip/body, vertical/horizontal)
- Covariates:
  - Segment duration
- Random effects:
  - Speaker
  - Previous segment, e.g. *sie sagt das* [ziː zaːkt das]
  - Internal segment, e.g. *sie sagt das* [ziː zaːkt das]

Variables: Classical Model

- Dependent variable:
  - Tongue sensor positions
- Main predictors:
  - Time (normalized)
  - Frequency (log)
  - Morphological conditions
  - Tongue types (e.g. tongue tip/body, vertical/horizontal)
- Covariates:
  - Segment duration
- Random effects:
  - Speaker
  - Previous segment, e.g. *sie sagt das* [ziː zaːkt das]
  - Internal segment, e.g. *sie sagt das* [ziː zaːkt das]
  - Next segment, e.g. *sie sagt das* [ziː zaːkt das]

Variables: Discriminative Model

Variables: Discriminative Model

- Dependent variable:
  - Tongue sensor positions
- Main predictors:
  - Time (normalized)
  - Frequency (log)
  - Morphological conditions
  - Tongue types (e.g. tongue tip/body, vertical/horizontal)
- Covariates:
  - Segment duration
- Random effects:
  - Speaker
  - Previous segment, e.g. *sie sagt das* [ziː zaːkt das]
  - Internal segment, e.g. *sie sagt das* [ziː zaːkt das]
  - Next segment, e.g. *sie sagt das* [ziː zaːkt das]

Variables: Discriminative Model

- Dependent variable:
  - Tongue sensor positions
- Main predictors:
  - Time (normalized)
  - Frequency (log)
  - Morphological conditions
  - Tongue types (e.g. tongue tip/body, vertical/horizontal)
- Covariates:
  - Segment duration
- Random effects:
  - Speaker
  - Previous segment, e.g. *sie sagt das* [ziː zaːkt das]
  - Internal segment, e.g. *sie sagt das* [ziː zaːkt das]
  - Next segment, e.g. *sie sagt das* [ziː zaːkt das]

Variables: Discriminative Model

- Dependent variable:
  - Tongue sensor positions
- Main predictors:
  - Time (normalized)
  -
  -
  - Tongue types (e.g. tongue tip/body, vertical/horizontal)
- Covariates:
  - Segment duration
- Random effects:
  - Speaker
  - Previous segment, e.g. *sie sagt das* [ziː zaːkt das]
  - Internal segment, e.g. *sie sagt das* [ziː zaːkt das]
  - Next segment, e.g. *sie sagt das* [ziː zaːkt das]

Variables: Discriminative Model

- Dependent variable:
  - Tongue sensor positions
- Main predictors:
  - Time (normalized)
  - Relative Functional Load

  - Tongue types (e.g. tongue tip/body, vertical/horizontal)
- Covariates:
  - Segment duration
- Random effects:
  - Speaker
  - Previous segment, e.g. *sie sagt das* [ziː zaːkt das]
  - Internal segment, e.g. *sie sagt das* [ziː zaːkt das]
  - Next segment, e.g. *sie sagt das* [ziː zaːkt das]

Model Comparison

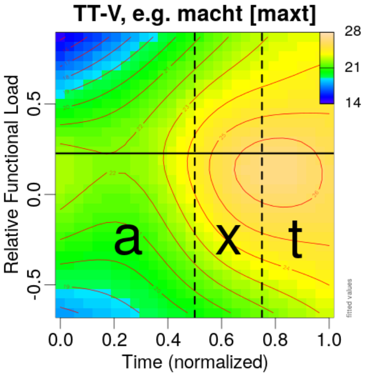
Model	Score	Edf	Diff.	Df	*p*-value
Freq.Morph	605825	78			
Rel.Func.Load	564386	30	-41439	48	

Model Comparison

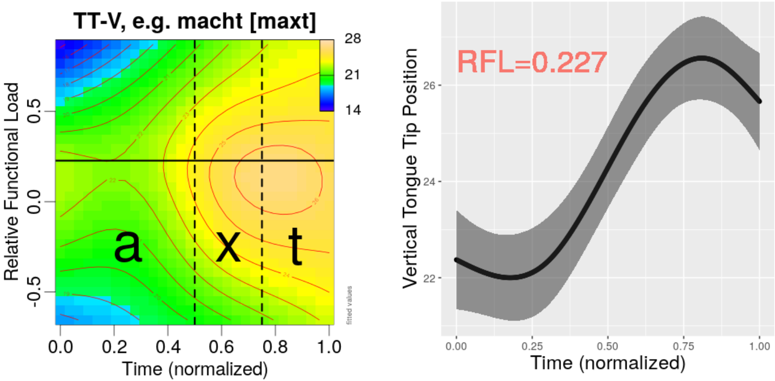
Model	Score	Edf	Diff.	Df	*p*-value
Freq.Morph	605825	78			
Rel.Func.Load	564386	30	-41439	48	

- Model with **Rel.Func.Load** is much **simpler AND better**.

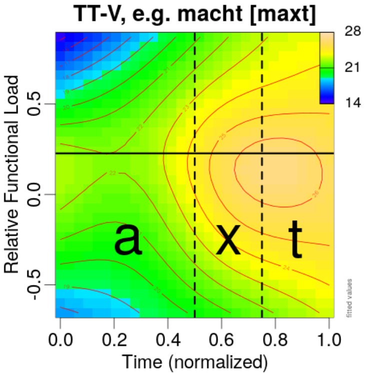
Tongue contours by time and Rel.Func.Load

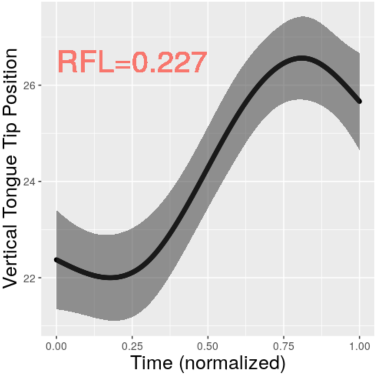


Tongue contours by time and Rel.Func.Load



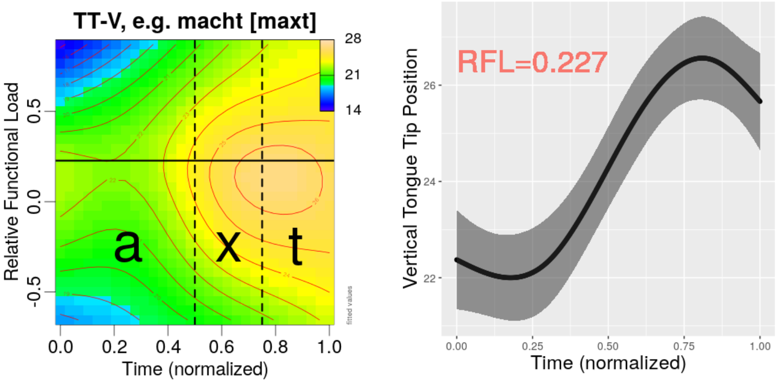
Tongue contours by time and Rel.Func.Load





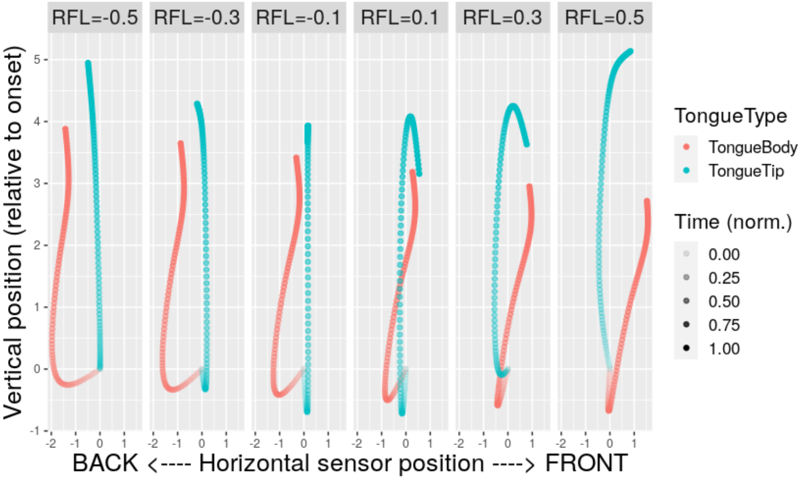
- Around the median of Rel.Func.Load

Tongue contours by time and Rel.Func.Load

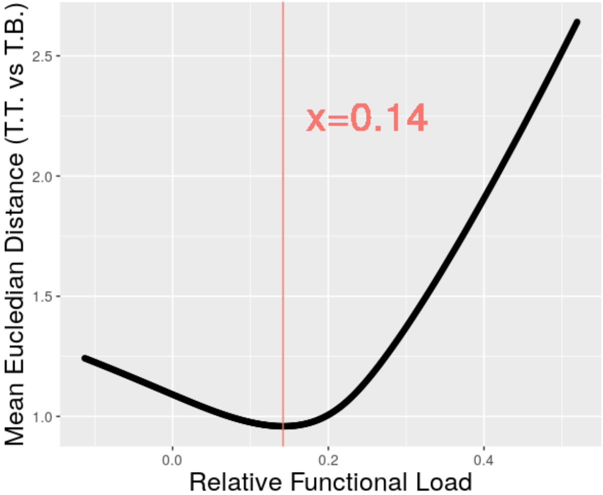


- Around the median of Rel.Func.Load
  - $\rightarrow$ the strongest coarticulation

Synchronization of Tongue Tip and Body

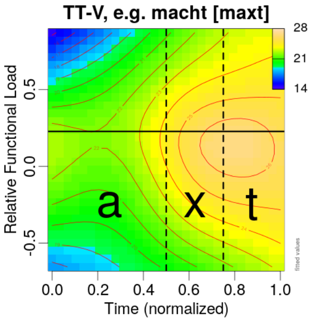


M.E.Distance between tongue tip and body trajectories

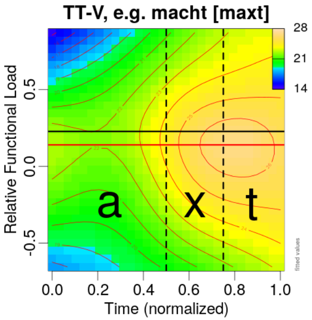


- Tongue tip and body are **synchronized the most** at **0.14**

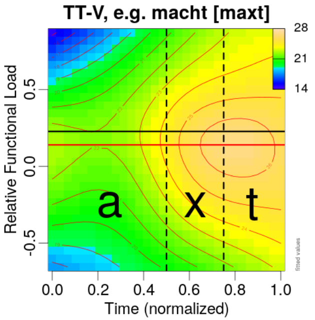
Tongue contours with the line of the most synchronization



Tongue contours with the line of the most synchronization

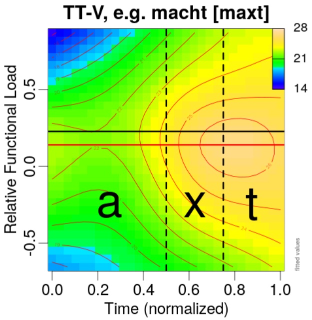


Tongue contours with the line of the most synchronization



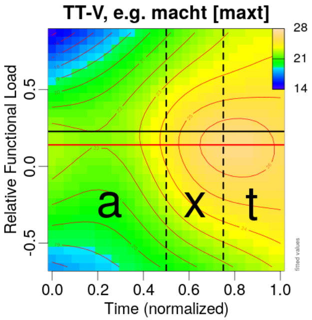
- Strongest coarticulation by the tongue tip

Tongue contours with the line of the most synchronization



- Strongest coarticulation by the tongue tip
- the greatest synchronization

Tongue contours with the line of the most synchronization

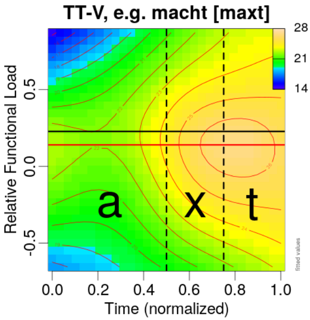


- Strongest coarticulation by the tongue tip
- the greatest synchronization

$\Downarrow$

- They coincide

Tongue contours with the line of the most synchronization



- Strongest coarticulation by the tongue tip
- the greatest synchronization

⇓

- They coincide

⇓

- **Optimal coarticulation**

Conclusion

- Tongue movements

Conclusion

- Tongue movements
  - $\rightarrow$ revealing and informative for the mental lexicon

Conclusion

- Tongue movements
  $\rightarrow$ revealing and informative for the mental lexicon
- EMA

Conclusion

- Tongue movements
  $\rightarrow$ revealing and informative for the mental lexicon
- EMA
  $\rightarrow$ interesting and powerful to study speech production

Conclusion

- Word frequency??
- Syllable frequency??

Conclusion

- Word frequency??
- Syllable frequency??

$\downarrow$

- Once semantics is considered straightforwardly,

Conclusion

- Word frequency??
- Syllable frequency??

$\downarrow$

- Once semantics is considered straightforwardly,
- Once the relation between forms and meanings is taken straightforwardly,

Conclusion

- Word frequency??
- Syllable frequency??

$\downarrow$

- Once semantics is considered straightforwardly,
- Once the relation between forms and meanings is taken straightforwardly,

$\Downarrow$

- It is irrelevant which frequency measure is “correct”.

Conclusion

- Word frequency??
- Syllable frequency??

$\downarrow$

- Once **semantics** is considered straightforwardly,
- Once the relation between forms and meanings is taken straightforwardly,

$\Downarrow$

- It is irrelevant which frequency measure is "correct".

Conclusion

- Word frequency??
- Syllable frequency??

$\downarrow$

- Once **semantics** is considered straightforwardly,
- Once the relation between forms and meanings is taken straightforwardly,

$\Downarrow$

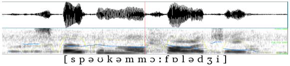
- It is irrelevant which frequency measure is “correct”.

$\Downarrow$

- **Semantics: all the way down**

Thank you very much!

Heinrich Heine Universität Düsseldorf



EBERHARD KARLS UNIVERSITÄT TÜBINGEN

This study is funded by the Deutsche Forschungsgemeinschaft (Research Unit FOR2373 'Spoken Morphology', Project 'The articulation of morphologically complex words', PL 151/7-1 and PL 151/8-1)

References I

Baayen, H., Chuang, Y.-Y., & Blevins, J. P. (2018). Inflectional morphology with linear mappings. *The Mental Lexicon*, *13*(2), 230–268. doi: 10.1075/ml.18010.baa

Baayen, R. H., Chuang, Y.-Y., Shafaei-Bajestan, E., & Blevins, J. P. (2019). The Discriminative Lexicon: A Unified Computational Model for the Lexicon and Lexical Processing in Comprehension and Production Grounded Not in (De)Composition but in Linear Discriminative Learning. *Complexity*, 1–39. doi: 10.1155/2019/4895891

Cho, T. (2001). Effects of Morpheme Boundaries on Intergestural Timing: Evidence from Korean. *Phonetica*, *58*, 129–162.

References II

Chuang, Y.-Y., Vollmer, M. L., Shafaei-Bajestan, E., Gahl, S., Hendrix, P., & 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

Hertrich, I., & Ackermann, H. (2000). Lipjaw and tonguejaw coordination during rate-controlled syllable repetitions. *The Journal of the Acoustical Society of America*, *107*(4), 2236–2247. doi: 10.1121/1.428504

Janssen, N., Bi, Y., & Caramazza, A. (2008). *A tale of two frequencies: Determining the speed of lexical access for Mandarin Chinese and English compounds* (Vol. 23) (No. 7-8). doi: 10.1080/01690960802250900

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–75.

References III

Tiede, M., Mooshammer, C., Goldstein, L., Shattuck-Hufnagel, S., & Perkell, J. S. (2011). Motor learning of articulator trajectories in the production of novel utterances. *Proceedings of the XVIIth International Congress of Phonetic Sciences*, 1986–1989.

Tomaschek, F., Arnold, D., Bröker, F., & Baayen, R. H. (2018). Lexical frequency co-determines the speed-curvature relation in articulation. *Journal of Phonetics*, *68*, 103–116.

Tomaschek, F., Plag, I., Ernestus, M., & 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. doi: 10.1017/S0022226719000203

Tomaschek, F., Tucker, B. V., & Baayen, R. H. (2019). How is anticipatory coarticulation of suffixes affected by lexical proficiency? *PsyArXiv (Preprent)*, 1–34.

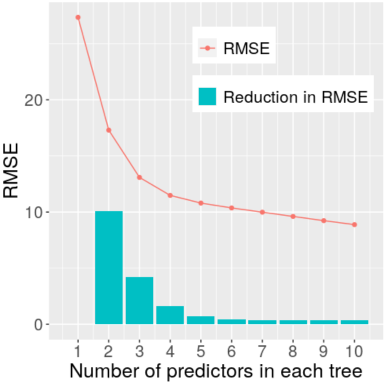
References IV

Tomaschek, F., Tucker, B. V., Fasiolo, M., & Baayen, R. H. (2018). Practice makes perfect: the consequences of lexical proficiency for articulation. *Linguistics Vanguard*, *4*.

Tomaschek, F., Wieling, M., Arnold, D., & Baayen, H. (2013). Word frequency, vowel length and vowel quality in speech production: An EMA study of the importance of experience. In *Proceedings of the 14th annual conference of the international speech communication association (interspeech 2013)* (pp. 1302–1306).

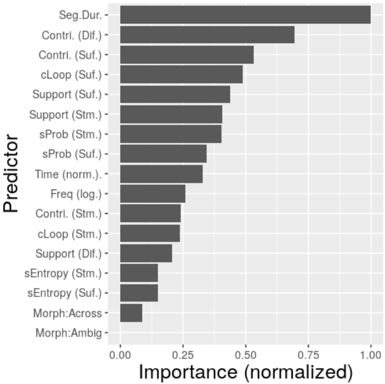
Appendix

Hyperparameter selectoin for Random Forest

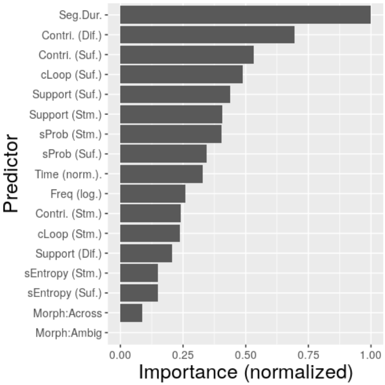


- 20 predictors in total
- Rule of thumb 1:
  - $\rightarrow$ 1/3 of the total number of predictors
  - $\rightarrow$ **6**
- Rule of thumb 2:
  - $\rightarrow$ Square root (rounded down) of the total number of predictors
  - $\rightarrow$ **4**
- Number of predictors $= 5$ is adopted in the present study.

Predictor selection by Random Forest

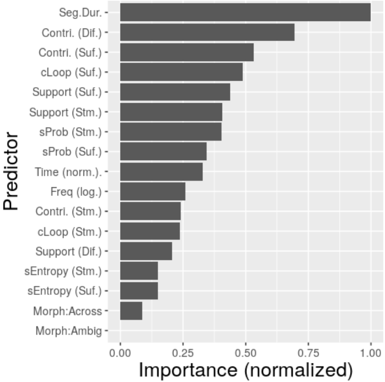


Predictor selection by Random Forest



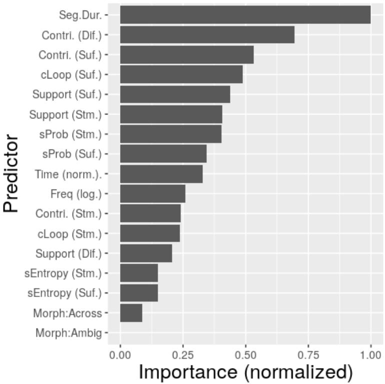
- Morphological status

Predictor selection by Random Forest



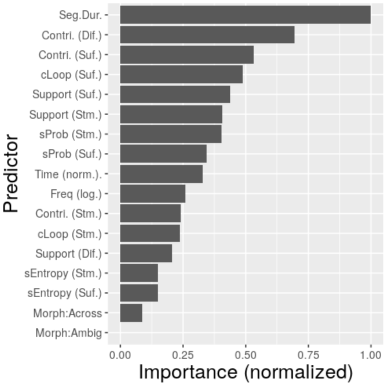
- Morphological status
  $\rightarrow$ not a strong predictor.

Predictor selection by Random Forest



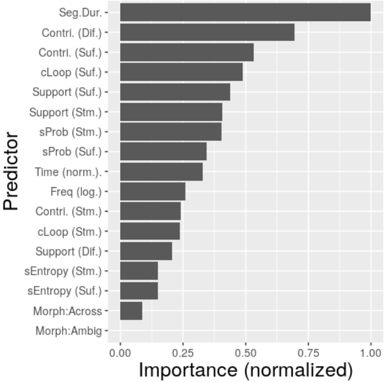
- Morphological status
  $\rightarrow$ not a strong predictor.
- Frequency

Predictor selection by Random Forest



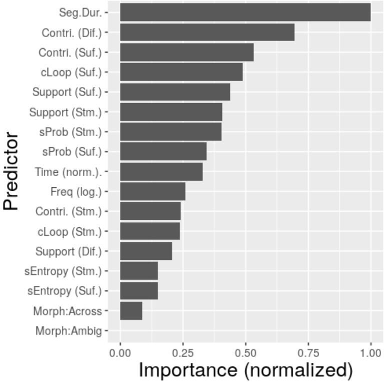
- Morphological status
  - $\rightarrow$ not a strong predictor.
- Frequency
  - $\rightarrow$ in the middle.

Predictor selection by Random Forest



- Morphological status
  - $\rightarrow$ not a strong predictor.
- Frequency
  - $\rightarrow$ in the middle.
- **Contri.(Dif.) = Relative Functional Load**

Predictor selection by Random Forest



- Morphological status
  - → not a strong predictor.
- Frequency
  - → in the middle.
- **Contri.(Dif.) = Relative Functional Load**
  - → The most effective among LDL-derived measures.

Training an LDL model

- An LDL model is trained on the target words and their inflectionally-related forms.

  e.g. *macht*, *machen*, *machst*, ...
- Simulated semantic vectors following H. Baayen, Chuang, and Blevins (2018).
- Accuracies:
  - Comprehension (forms →meanings)

    → 91.7%
  - Production (meanings →forms)

    → 90.8%

Mapping between forms and meanings

$$C = \begin{array}{c} \\ hand \\ and \\ band \\ \ldots \end{array} \begin{array}{c} \begin{array}{ccccc} \#ha & han & \#an & and & \ldots \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 1 & \ldots \\ 0 & 0 & 1 & 1 & \ldots \\ 0 & 0 & 0 & 1 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \end{pmatrix} \end{array}$$

$$S = \begin{array}{c} \\ hand \\ and \\ band \\ \ldots \end{array} \begin{array}{c} \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & \ldots \end{array} \\ \begin{pmatrix} 0.989 & 0.915 & 0.232 & 0.190 & \ldots \\ 0.004 & 0.101 & 0.892 & 0.380 & \ldots \\ 0.643 & 0.004 & 0.401 & 0.899 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \end{pmatrix} \end{array}$$

Mapping between forms and meanings

$$C = \begin{array}{c c} & \begin{array}{ccccc} \#ha & han & \#an & and & \dots \end{array} \\ \begin{array}{c} hand \\ and \\ band \\ \dots \end{array} & \begin{pmatrix} 1 & 1 & 0 & 1 & \dots \\ 0 & 0 & 1 & 1 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} \end{array}$$

$$S = \begin{array}{c c} & \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & \dots \end{array} \\ \begin{array}{c} hand \\ and \\ band \\ \dots \end{array} & \begin{pmatrix} 0.989 & 0.915 & 0.232 & 0.190 & \dots \\ 0.004 & 0.101 & 0.892 & 0.380 & \dots \\ 0.643 & 0.004 & 0.401 & 0.899 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} \end{array}$$

$$C\color{red}{F}\color{black} = S$$

Conceptual understanding of transformation matrices

$$\begin{bmatrix} 3 \\ 1 \end{bmatrix} \cdot ? = \begin{bmatrix} 9 \\ 3 \end{bmatrix}$$

Conceptual understanding of transformation matrices

$$\begin{bmatrix}3\\1\end{bmatrix} \cdot {\color{red}3} = \begin{bmatrix}9\\3\end{bmatrix}$$

Conceptual understanding of transformation matrices

$$\begin{bmatrix} 3 \\ 1 \end{bmatrix} \cdot ? = \begin{bmatrix} 10 \\ 3 \end{bmatrix}$$

Conceptual understanding of transformation matrices

$$\begin{bmatrix}3\\1\end{bmatrix} \cdot {\color{red}3} = \begin{bmatrix}9\\3\end{bmatrix} \neq \begin{bmatrix}10\\3\end{bmatrix}$$

Conceptual understanding of transformation matrices

$$\begin{bmatrix} 3 \\ 1 \end{bmatrix} \cdot {\color{red}3} = \begin{bmatrix} 9 \\ 3 \end{bmatrix} \neq \begin{bmatrix} 10 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} 3 \\ 1 \end{bmatrix} \cdot {\color{blue}5} = \begin{bmatrix} 15 \\ 5 \end{bmatrix} \neq \begin{bmatrix} 10 \\ 3 \end{bmatrix}$$

Conceptual understanding of transformation matrices

$$\begin{bmatrix} 3 \\ 1 \end{bmatrix} \cdot {\color{red}3} = \begin{bmatrix} 9 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} 9 \\ 3 \end{bmatrix} \cdot {\color{blue}\frac{1}{3}} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

Conceptual understanding of transformation matrices

$$\begin{bmatrix} 3 \\ 1 \end{bmatrix} \cdot {\color{red}3} = \begin{bmatrix} 9 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} 9 \\ 3 \end{bmatrix} \cdot {\color{blue}\frac{1}{3}} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$C{\color{red}F} = S$$

$$S{\color{blue}G} = C$$

Mapping between C and S via a weight matrix F

$$C = \begin{array}{c c} & \begin{array}{ccccc} \#ha & han & \#an & and & \dots \end{array} \\ \begin{array}{c} hand \\ and \\ band \\ \dots \end{array} & \begin{pmatrix} 1 & 1 & 0 & 1 & \dots \\ 0 & 0 & 1 & 1 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} \end{array}$$

$$S = \begin{array}{c c} & \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & \dots \end{array} \\ \begin{array}{c} hand \\ and \\ band \\ \dots \end{array} & \begin{pmatrix} 0.989 & 0.915 & 0.232 & 0.190 & \dots \\ 0.004 & 0.101 & 0.892 & 0.380 & \dots \\ 0.643 & 0.004 & 0.401 & 0.899 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} \end{array}$$

$$F = \begin{array}{c c} & \begin{array}{ccccc} S_1 & S_2 & S_3 & S_4 & \dots \end{array} \\ \begin{array}{c} \#ha \\ han \\ \#an \\ \dots \end{array} & \begin{pmatrix} 0.739 & 0.332 & 0.392 & 0.293 & \dots \\ 0.231 & 0.384 & 0.904 & 0.224 & \dots \\ 0.610 & 0.092 & 0.119 & 0.028 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} \end{array}$$

Mapping by transformation matrices

$$CF = S$$

$$SG = C$$

- Transformation matrix F/G is estimated, given C and S.
- Conceptually, F/G is the learned language processing system.
  - It can predict semantics, based on forms (e.g. tri-phones)
    $\rightarrow$ $CF = \hat{S}$
  - It can predict forms, based on semantics.
    $\rightarrow$ $SG = \hat{C}$

Model structure

```
SenTT.Z ~ TongueMorph
        + te(Time, Freq)
        + te(Time, Freq, by=TongueMorph)
        + s(SegmentDuration)
        + s(Speaker, bs='re')
        + s(PrevSeg, bs='re')
        + s(IntSeg,  bs='re')
        + s(NextSeg, bs='re')
```

Result: Model Summary (Parametric terms)

(A. ParametricTerms)	Estimate	Std.Error	t value	p-value
Intercept	-68.125	1.586	-42.947	0.000
TongueMorph=TBX.1	-0.216	0.649	-0.333	0.739
TongueMorph=TBX.2	-0.348	0.626	-0.555	0.579
TongueMorph=TBZ.0	79.232	0.479	165.294	0.000
TongueMorph=TBZ.1	79.762	0.645	123.573	0.000
TongueMorph=TBZ.2	80.247	0.628	127.794	0.000
TongueMorph=TTX.0	21.997	0.479	45.914	0.000
TongueMorph=TTX.1	24.186	0.646	37.440	0.000
TongueMorph=TTX.2	22.851	0.629	36.356	0.000
TongueMorph=TTZ.0	78.034	0.479	162.988	0.000
TongueMorph=TTZ.1	78.113	0.643	121.494	0.000
TongueMorph=TTZ.2	78.856	0.630	125.090	0.000

Result: Model Summary (Smooth Terms) (Main Predictors)

(B. SmoothTerms)	edf	Ref.df	F	p-value
te(Time,Freq):TM=TBX.0	7.055	8.086	3.743	0.000
te(Time,Freq):TM=TBX.1	5.162	5.991	2.792	0.011
te(Time,Freq):TM=TBX.2	9.806	11.300	6.872	0.000
te(Time,Freq):TM=TBZ.0	5.058	5.946	7.817	0.000
te(Time,Freq):TM=TBZ.1	4.355	5.000	5.496	0.000
te(Time,Freq):TM=TBZ.2	13.503	15.859	7.210	0.000
te(Time,Freq):TM=TTX.0	4.789	5.594	5.880	0.000
te(Time,Freq):TM=TTX.1	3.120	3.232	9.722	0.000
te(Time,Freq):TM=TTX.2	11.728	13.598	8.529	0.000
te(Time,Freq):TM=TTZ.0	5.634	6.548	19.269	0.000
te(Time,Freq):TM=TTZ.1	6.483	7.316	11.036	0.000
te(Time,Freq):TM=TTZ.2	13.563	15.834	12.682	0.000

Result: Model Summary (Smooth Terms) (Covariates & REs)

(B. SmoothTerms)	edf	Ref.df	F	p-value
s(SegmentDuration)	1.045	1.088	1.621	0.180
s(Speaker)	35.863	36.000	1201.290	0.000
s(PreviousSegment)	16.637	19.000	150.671	0.000
s(InternalSegment)	5.647	10.000	83.385	0.016
s(NextSegment)	34.284	63.000	36.269	0.000

Model structure (LDL-measure model)

```
SenTT.Z ~ TongueType
        + te(Time, RelFuncLoad)
        + te(Time, RelFuncLoad, by=TongueType)
        + s(SegmentDuration)
        + s(Speaker, bs='re')
        + s(PrevSeg, bs='re')
        + s(IntSeg,  bs='re')
        + s(NextSeg, bs='re')
```

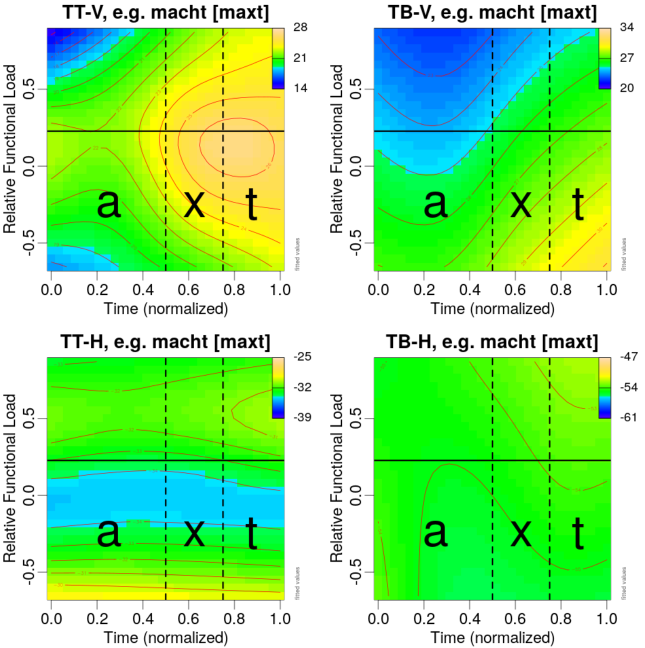
Result: Model Summary (Parametric terms)

(A. ParametricTerms)	Estimate	Std.Error	t value	p-value
Intercept	-67.790	1.510	-44.898	0.000
TongueType=TBZ	79.970	0.174	458.493	0.000
TongueType=TTX	21.424	0.174	123.433	0.000
TongueType=TTZ	78.278	0.175	448.457	0.000

Result: Model Summary (Smooth Terms)

(B. SmoothTerms)	edf	Ref.df	*F*	*p*-value
te(Time,RelFuncLoad):TBX	6.400	7.616	3.744	0.000
te(Time,RelFuncLoad):TBZ	6.491	7.494	19.554	0.000
te(Time,RelFuncLoad):TTX	10.538	12.938	6.126	0.000
te(Time,RelFuncLoad):TTZ	14.062	17.307	16.840	0.000
s(SegmentDuration)	1.775	2.236	2.474	0.099
s(Speaker)	35.841	36.000	1440.440	0.000
s(PreviousSegment)	16.313	19.000	202.549	0.000
s(InternalSegment)	4.465	10.000	41.164	0.345
s(NextSegment)	37.535	63.000	53.698	0.000

Tongue contours by time and Rel.Func.Load all 4 tongue types



Tongue contours and synchronization all 4 tongue types