On Functional Load and its relation to the actuation problem

Andrea Ceolin - University of Pennsylvania
NWAV48 - October 10th, 2019

Section 1

The Functional Load Hypothesis

- The likelihood of a merger between two phonemes is inversely proportional to the number of word pairs distinguished by those two phonemes (Mathesius 1929, Jakobson 1931, Martinet 1955)

- 'Martinet has cogently and persistently argued that functional load has rich yet unexplored possibilities for the linguist who attempts to plumb the causality of sound change' (King 1967:833)

The Functional Load Hypothesis

- Very few empirical investigations, inconsistent results.

- Reports of sound changes that violate Functional Load: King (1967), Surendran and Niyogi (2006), Sampson (2015), Cohen Priva (2017)

- Reports of sound changes that are consistent with Functional Load: Gurevich (2004), Bouchard-Côté et al. (2013), Wedel et al. (2013), Eychenne and Jang (2018), Babinsky and Bowern (2018)

Wedel et al. (2013)

- 'Big data' approach: they examine a large collection of mergers reported in English Received Pronunciation (RP), American English, German, Dutch, French, Spanish, Slovak, Korean and Cantonese

- The mergers are coded as a dependent variable in a mixed effects logistic regression model

- They find a significant effect of Functional Load measured as the number of **Minimal Pairs** associated with a contrast

- **Frequency** and **Entropy** are not (or barely) significant

Wedel et al. (2013)

- Things which are great about this paper:
    - Focus on ongoing mergers: a lot of documentation
    - Use of minimal pairs
- Things which can be improved about this paper:
    - Phonetic control (not today)
    - CELEX (today)
    - Mergers (today)
- In this talk, we attempt a replication.

Section 2

CHILDES Data

- CELEX is not a realistic corpus of child directed utterances
- We collected CHILDES data from English, Dutch, German and Spanish and transcribed them to phonological representation
- Filters: lemma, more than one occurrence in total across different corpora (list in the Appendix), real words
- Total: 4925 for English, 4273 for Dutch, 3066 for German, 2911 for Spanish

Mergers

English RP	aɪ-ɔɪ, ʊə-ɔː
	ɪə-ɛə, ɜː-ɛə (RP mergers? cf. Sampson, *to appear*)
	θ-f, ð-v
	θ-t, ð-d (exogenous change, cf. Honeybone 2016)
	θ-s, ð-z (not reported in Wells 1982)
American English	ɑ-ɔ, i-ɪ (_l), ʊ-u (_l), ʊ-oʊ (_l), ʌ-ɑ (_l), ʊ-ʌ(_l), ɪ-ɛ (_n), t-d (V́_V)
	ɑ-ɔ (_r) (interaction with the *horse/hoarse* merger)
German	e-ɛː
Dutch	s-z, f-v, χ-ɣ
Spanish	θ-s, ʎ-j (cf. not a recent merger, Harris 1967)

Table 1: The mergers investigated in this work. These mergers are a subset of those reported in Wedel et al. (2013). They were selected based on literature references (Harris 1967, Wells 1982, Wiese 2000, Kissine et al. 2003, Blevins 2004, Labov et al. 2007, Honeybone 2016)

Model

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-2.29663	0.41135	-5.583	2.36e-08 ***
MinPairs	-0.12151	0.05914	-2.055	0.0399*
FreqMax	0.09287	0.27464	0.338	0.7753
FreqMin	-0.03393	0.34601	-0.098	0.9219

Table 2: Mixed-effects Logistic Regression, with Merger Set as Random effect

Positive test for Functional Load. What next?

Consistent with Functional Load

English RP	aɪ-ɔɪ	*by-boy, buy-boy, bye-boy, oil-isle, ...* [+3]
	ʊə-ɔː	
	θ-f	*half-hearth, three-free, thin-fin, thought-fort*
	ð-v	*than-van*
American English	ɑ-ɔ	
	ʌ-ɑ (_l)	
	ʊ-ʌ(_l)	
German	e-ɛː	
Dutch	s-z	
	f-v	
	χ-ɣ	
Spanish	ʎ-j	

Confusion Indexes do not explain the mergers

Language	Merger	MinPairs	Confusion
English RP	θ-f	4	16.1
English RP	ð-v	1	10.9
English RP	f-v	17	5
English RP	s-z	15	4.9
English RP	p-b	66	2.8
English RP	t-d	80	2
English RP	p-f	52	12.3
Dutch	f-v	0	18.98
Dutch	s-z	1	13.49
Dutch	p-b	43	13.22
Dutch	t-d	21	18.38

Table 3: Confusion indexes from Smits et al. (2003), Weber and Smits (2003)

Inconsistent with Functional Load

English RP	ɪə-ɛə	*beer-bear, deer-dare, fear-fair, cheer-chair, ...* [+12]
	ɜː-ɛə	*burly-barely, furry-fairy, fur-fair, her-hair,...* [+7]
American English	i-ɪ (_l)	*feel-fill, heel-hill, meal-mill, peel-pill,...* [+3]
	ʊ-u (_l)	*full-fool, pull-pool*
	ʊ-oʊ (_l)	*bull-bowl, full-foal, pull-pole*
	ɪ-ɛ (_n)	*din-den, mini-many, pin-pen, since-sense,...* [+2]
	t-d (V́_V)	*writer-rider, petal-pedal*

Most of the other vowel contrasts in these contexts yield zero
minimal pairs!

Section 3

Inconsistent with Functional Load

Sound changes violating Functional Load

English RP	ɪə-ɛə	
	ɜː-ɛə	
American English	i-ɪ (_l), ʊ-u (_l), ʊ-oʊ (_l)	
	ɪ-ɛ (_n)	
	t-d (V́_V)	
Hong Kong Cantonese	n-l (#_)	Surendran and Niyogi (2006)
American English	t-deletion (_#)	Cohen Priva (2017)
Spanish	s-deletion (_#)	Cohen Priva (2017)
Middle Chinese	[p-/t-/k-deletion] (_#)	Sampson (to appear)
Middle Chinese	m-n (_#)	Sampson (to appear)
Middle Chinese	palatalization (_[+front])	Sampson (to appear)

Inconsistent with Functional Load

Sound changes violating Functional Load

English RP	ɪə-ɛə	
	ɜː-ɛə	
American English	i-ɪ (_l), ʊ-u (_l), ʊ-oʊ (_l)	
	ɪ-ɛ (_n)	
	t-d (V́_V)	
Hong Kong Cantonese	n-l (#_)	Surendran and Niyogi (2006)
American English	t-deletion (_#)	Cohen Priva (2017)
Spanish	s-deletion (_#)	Cohen Priva (2017)
Middle Chinese	[p-/t-/k-deletion] (_#)	Sampson (to appear)
Middle Chinese	m-n (_#)	Sampson (to appear)
Middle Chinese	palatalization (_[+front])	Sampson (to appear)

These are conditioned mergers! (primary splits, Hoenigswald 1965)

Acquisition

- Unconditioned mergers result from the failure in acquiring a categorical contrast. Categories are acquired in the first months of life, so these changes must happen early (Werker and Tees 1984, Kuhl et al. 1992)

- Lexical contrast is a plausible cue to identify categories (Cui 2019)

Acquisition

- Conditioned mergers, on the other hand, have to do with the reanalysis of phonological rules
- They (can) apply when the categories are in place
- At this stage, other things matter (e.g., rule productivity, Richter 2017)

Summary

- Assessing the plausibility of Functional Load empirically is a hard task

- As scientists, we are not happy with statistical tendencies, but we want to have hypotheses that come with predictions (King 1967, Sampson 2015)

- I propose that Functional Load might be a consequence of the way category learning works in natural languages

- Sound changes that do not necessarily involve category loss (e.g. loss of allophonic rule, lenition) do not relate to Functional Load

Thanks!

Thanks to Ryan Budnick, Spencer Caplan, Aletheia Cui, Jordan Kodner, Caitlin Richter, Ollie Sayeed, David Wilson, Bob Berwick, Robin Clark, Rolf Noyer, Don Ringe, Gareth Roberts, Geoff Sampson, Meredith Tamminga, Charles Yang.
Thanks to NWAV48 organizers!

CHILDES: Corpora

English: *Providence*

Dutch: *Groningen, Utrecht, VanKampen*

German: *Rigol, Miller*

Spanish: *Ornat, Vila, Nieva, OreaPine*

Mergers in American English

CMU dictionary

far —— for   four   far   for — four

Spanish (Harris 1967)