The actuation of sound change: confusion or functional considerations?
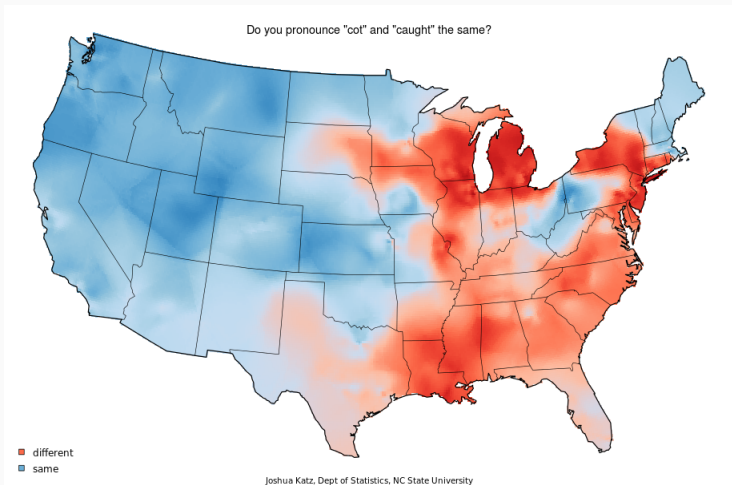
Andrea Ceolin
University of Pennsylvania

Figure 1: The caught/cot merger. Why did it happen?

Solutions

Solution n.1: these sounds are very easy to confuse.

Intended	Perceived	p
/ɔ/	/ɑ/	**0.138**
/æ/	/ɛ/	0.056
/ʌ/	/ɑ/	0.037
/ɛ/	/æ/	0.037

Table 1: Vowel confusability indexes from Hillenbrand et al. (1995)

Solutions

Solution n.2: these sounds are not contrastive.

Contrast	Minimal Pairs	Words
/ɔ/-/ɑ/	**0**	
/æ/-/ɛ/	11	bad/bed, bat/bet, gas/guess, had/head...
/ʌ/-/ɑ/	2	but/bought, cut/cought

Table 2: Minimal pairs from a CHILDES wordlist

FUNCTIONAL LOAD

The likelihood of a merger between two phonemes is inversely proportional to the number of pairs distinguished by such two phonemes (Mathesius 1929, Martinet 1955, Jakobson 1975)

The Functional Load Hypothesis

Very few empirical investigations (King 1967, Surendran and Niyogi 2003, Wedel et al. 2013, Eychenne and Jang 2018)

Wedel et al. (2013) has been influential among linguists (Sóskuthy 2013, Kiparsky 2016, Babinsky and Bowern 2018)

'Big data' approach: they examine a large collection of mergers reported in British and American English, German, Dutch, French, Spanish, Slovak, Korean and Cantonese

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

CORPORA Minimal pairs are counted using the whole CELEX corpora. Not plausible from an acquisition viewpoint.

PHONETICS Merged pairs are evaluated against unmerged pairs with equal featural distance: e.g., /t/-/d/ is evaluated against /t/-/k/

MISTAKES Appendix of the paper reveals some mistakes

CORPORA Use CHILDES data or, in alternative, frequency threshold on CELEX

PHONETICS Add a confusability index as predictor (Wang and Bilger 1967, Hillenbrand et al. 1995, Weber and Smits 2003, Smits et al. 2003, Jouvet et al. 2015)

MISTAKES Cross-check every merger with the literature (Wells 1982, Labov et al. 2006, Wiese 2000, Kissine et al. 2003)

Data (roughly 50% of the dataset)

British English		American English		German
θ-f		ɑ-ɔ	LOT-THOUGHT	
ð-v		*æ-ɑ		e-ɛ:
θ-t		*ɔɪ-ɚ		
ð-d				
*θ-s		_l		
*ð-z		i-ɪ	HILL-HELL	
		ʊ-u	PULL-POOL	
aɪ-ɔɪ	PRICE-CHOICE	ʊ-oʊ	BULL-BOWL	Dutch
ʊə-ɔ:	CURE-THOUGHT	ʌ-ɑ	HULL-HALL (ʌ-ɔ)	
ɪə-ɛə	NEAR-SQUARE	ʊ-ʌ	BULL-HULL	s-z
ɜ:-ɛə	NURSE-SQUARE			f-v
		_n		χ-ɣ
		ɪ-ɛ	PIN-PEN	

Model

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.48466	0.57321	-2.590	0.00959*
minPairs	-0.35859	0.17520	-2.047	**0.04068***
phonemeFreq	0.37262	0.36551	1.019	0.30798
NoMP	-0.70144	0.70198	-0.999	0.31769
phonFreq by NoMP	-0.09863	0.61097	-0.161	0.87175

Table 3: Mixed-effect Logistic Regression, with Merger Set as Random effect.
AIC 130.0

Model

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-2.38577	0.65193	-3.660	0.000253*
minPairs	-0.35312	0.16343	-2.161	**0.030720***
phonemeFreq	0.18824	0.46323	0.406	0.684473
NoMP	-0.64017	0.72233	-0.886	0.375479
confusability	0.06586	0.01740	3.784	**0.000154***
phonFreq by NoMP	-0.35478	0.69297	-0.512	0.608677

Table 4: Mixed-effect Logistic Regression, with Merger Set as Random effect. AIC 119.8

Model

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-2.97224	0.52261	-5.687	<0.000001*
minPairs	-0.15511	0.13727	-1.130	**0.2585**
confusability	0.06617	0.02403	2.754	**0.0059***

Table 5: Mixed-effect Logistic Regression, with Merger Set as Random effect, for Vowels

Minimal Pairs

British English (average: 5.6)

ai-oi 4 (*tie/toy*, *bye/boy*, *by/boy*, *buy/boy*)

ɜː-ɛə 3 (*her/hair*, *where/wear*, *where/were*)

American English (before /l/) (average: 0.5)

i-ɪ 2 (*feel/fill*, *wheel/will*)

German (average: 5.2)

e-ɛː 5 (*fällen/fehlen*, *stehlen/stellen*, *her/herr*, *Heer/herr*, *Weg/weg*)

Model

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-2.97430	0.84338	-3.527	0.000421*
minPairs	-0.37884	0.18045	-2.099	**0.035782***
confusability	0.07983	0.03012	2.651	**0.008028***

Table 6: Mixed-effect Logistic Regression, with Merger Set as Random effect, for Consonants

The Functional Load Hypothesis

Language	Merger	MP	CI (sum)	CI (max)
British English	θ-f	1	32.3	22.1
British English	ð-v	0	21.9	11.7
British English	θ-t	4	17.1	11.7
British English	ð-d	2	23.5	13.3
British English	θ-s	4	9.8	9.4
British English	ð-z	0	20.4	15.4
British English	θ-ð	0	20.4	11.0
British English	p-f	**6**	24.6	16.5
British English	b-v	0	20.8	10.6
Dutch	f-v	1	58.14	29.16
Dutch	s-z	0	39.96	20.44
Dutch	p-b	8	49.25	37.96
Dutch	t-d	**15**	90.34	85.40

Table 7: Consonant mergers

Conclusion

1. This problem is hard. No solution so far. In some case there is clearly no Functional Load effect, in other cases there is...

2. This is an important empirical problem: it has a lot to say about what phonemes are, how they are acquired, and how they can change.

3. There are only few people working on this problem :(Join us!

Thank you!

Thanks to Ryan Budnick, Spencer Caplan, Aletheia Cui, Jordan Kodner, Caitlin Richter, Ollie Sayeed, David Wilson, Bob Berwick, Robin Clark, Rolf Noyer, Don Ringe, Meredith Tamminga, Charles Yang

Questions?