Examining Extragrammatical Effects on English Auxiliary Contraction

Constantine Lignos and Laurel MacKenzie

Penn University of Pennsylvania

MANCHESTER 1824 The University of Manchester

Overview

- Multiple factors condition contraction (e.g., /ɪz/ ~ /z/), including subject length (MacKenzie 2013) and predictability of auxiliary (Frank and Jaeger 2008)
- What is their relative contribution?
- What causes these effects?

Auxiliary contraction

- We examine contraction (use of underlying short form) of three English auxiliaries after non-pronoun subjects
- Auxiliaries surface in a number of phonological shapes:

is	[ɪz]	[əz]		[s], [z]
has	[hæz]	[həz]	[əz]	[s], [z]
will	[wɪl]	[wəl]	[əl]	[l]

- Underlyingly: bipartite allomorphic alternation:

	'long'		'short'
is	/ɪz/	~	/z/
has	/hæz/	~	/z/
will	/wɪl/	~	/l/

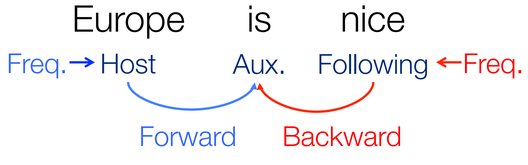
(Kaisse 1983, MacKenzie 2013)

- Taking into account existence of later processes (e.g. reduction, deletion), surface forms can be reclassified:

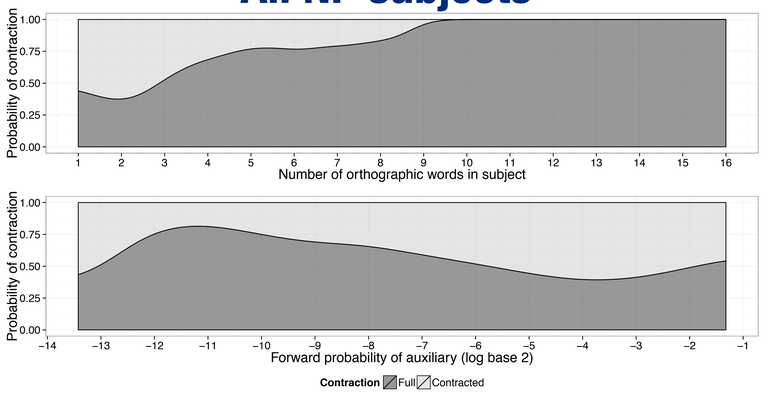
	Underlying long		Underlying short	
is	[ɪz]	[əz]		[s], [z]
has	[hæz]	[həz]	[əz]	[s], [z]
will	[wɪl]	[wəl]	[əl]	[l]

Data sources

- Switchboard, Fisher, Phila. Neighborhood Corpus
- 1092 tokens in contractible contexts with valid predictability information, coded auditorily
- Predictors:
  - Subject size (multiple measures)
  - Frequency, predictability:



All NP subjects



Modeling

- Mixed effects logistic regression using lme4.0
- Control fixed effects:
  - Subject-based: age***, gender*, education level***
  - Token-based: corpus, speaking rate, aux. identity, preceding segment (C or V) and stress
- Fixed effects of interest:
  - Subject size, frequency of host and following word (log), forward and backward probability of aux. (log)
- Random intercepts: preceding and following word, speaker dialect region, speaker identity

Results

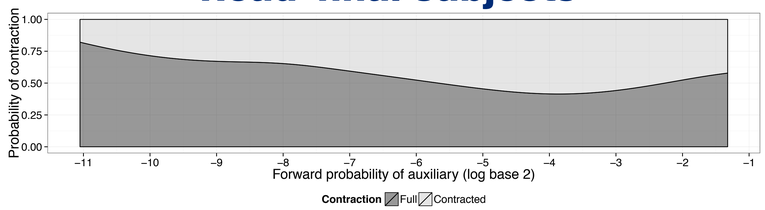
- N. words and forward prob. are most reliable:

Predictor	Estimate	Std. Error	P(>|z|)	
Number of words	-0.440	0.074	3.3E-09	***
Host freq.	0.063	0.036	0.078	.
Following word freq.	0.070	0.033	0.036	*
Forward prob.	0.242	0.060	5.7E-05	***
Backward prob.	0.098	0.059	0.095	.

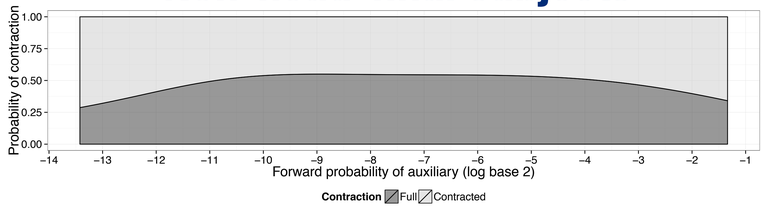
(Significance codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1)

- N. words has larger effect than forward prob. over useful range of predictor:
  - *N. words*: *Mary* (1) vs. *the county that I live in now* (7), increase of 6 words →14.1x less likely
  - *Forward prob.*: *Communism is* (.364) vs. *work is* (.00787), 46x (2^5.5) decrease → 3.8x less likely

Head-final subjects



Non-head-final subjects



Examining forward probability

- Lowest forward prob. items are often non-head final forms: *the guy next to you is* (*you* → *is* improbable)
- Model interaction of head-finality
- Forward prob. only significant (p = 4.7E-05) for head-final subjects, non-head-final n.s. (p = 0.086)

Examining subject size

- Many measurements: duration, n. words, n. prosodic/ function words, n. syllables, embedding depth
- Duration provides better fit than n. words, but n. words is better than all other measures

Conclusions

- Size and predictability matter, but size has larger effect
- Predictability effects only found for head-final NPs, further exploration needed to identify mechanism
- Connections to be made with other length-conditioned variable phenomena (e.g., heavy NP shift, dative alternation) and models of production/variation

References

Frank, Austin, and T. Florian Jaeger. 2008. Speaking rationally: Uniform information density as an optimal strategy for language production. In *The 30th Annual Meeting of the Cognitive Science Society* (CogSci08), 939–944.

Kaisse, Ellen M. 1983. The syntax of auxiliary reduction in English. *Language* 59:93–122.

MacKenzie, Laurel. 2013. Variation in English auxiliary realization: A new take on contraction. *Language Variation and Change* 25:17–41.