Filling in the blanks in morphological productivity: a word-completion task

Kyle Mahowald, Timothy J. O'Donnell, and Joshua B. Tenenbaum

Motivation

Are humans systematically biased in judgments of language frequency?

-In a classic study, (Tversky & Kahneman, 1973, 1983), participants were asked to judge how many words in a hypothetical text of 2,000 English words would fit the pattern _ _ _ _ _ n _. The median guess was *higher* for _ _ _ _ i n g than for _ _ _ _ _ n _ , despite this being a logical impossibility.

-Tversky and Kahneman interpret this result as a failure of rational judgment, and attribute it the greater *availability* of words ending in *-ing*.

-We believe this greater availability follows naturally from morphological productivity, and that people's overestimation for **productive** morphemes follows rationally from the ability of these morphemes to form new words.

-We test this hypothesis by extending Tversky and Kahneman's work using a variety of English derivational suffixes of varying levels of productivity. We derive productivity estimates for various suffixes and predict that more productive suffixes (e.g., -ness) are more likely to have their frequencies overestimated than lower productivity suffixes (e.g., -ity).

How many words fit this pattern?

_ _ _ _ _ n _

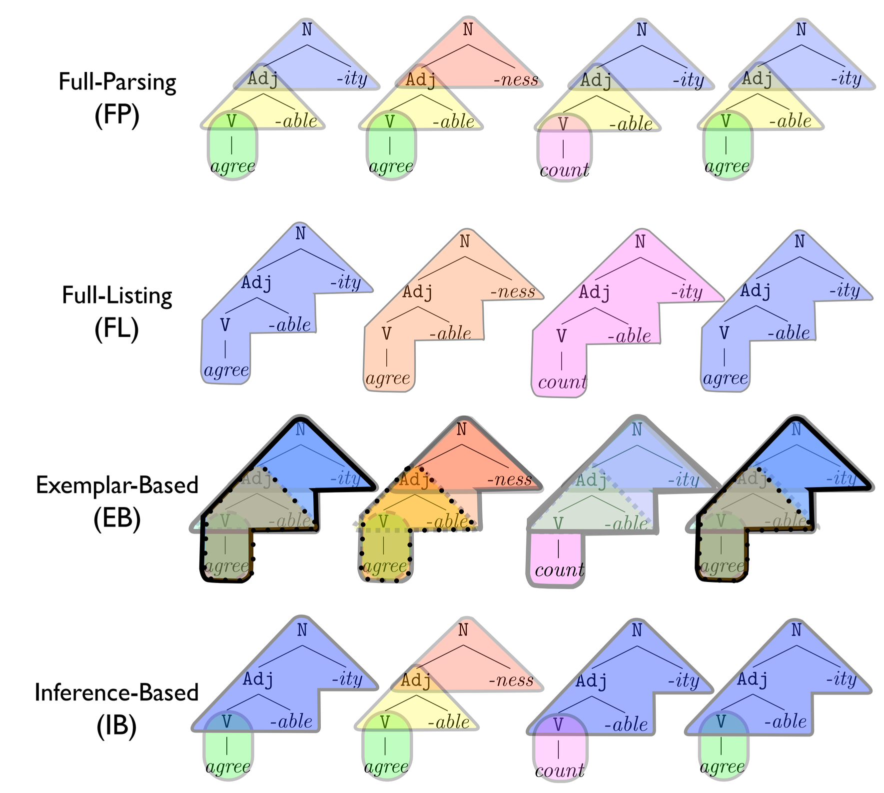
_ _ _ _ i n g

Morphological Productivity

-The ability to give rise to novel forms in language is known as linguistic *productivity* (Bauer 2001). This phenomenon is evident in derivational morphology.

	Suffix	examples
Productive	-ness	happiness, psycholinguistic-ness, coolness
Semi-productive	-ity	sparsity, scarcity, coolity*
Unproductive	-th	warmth, coolth*

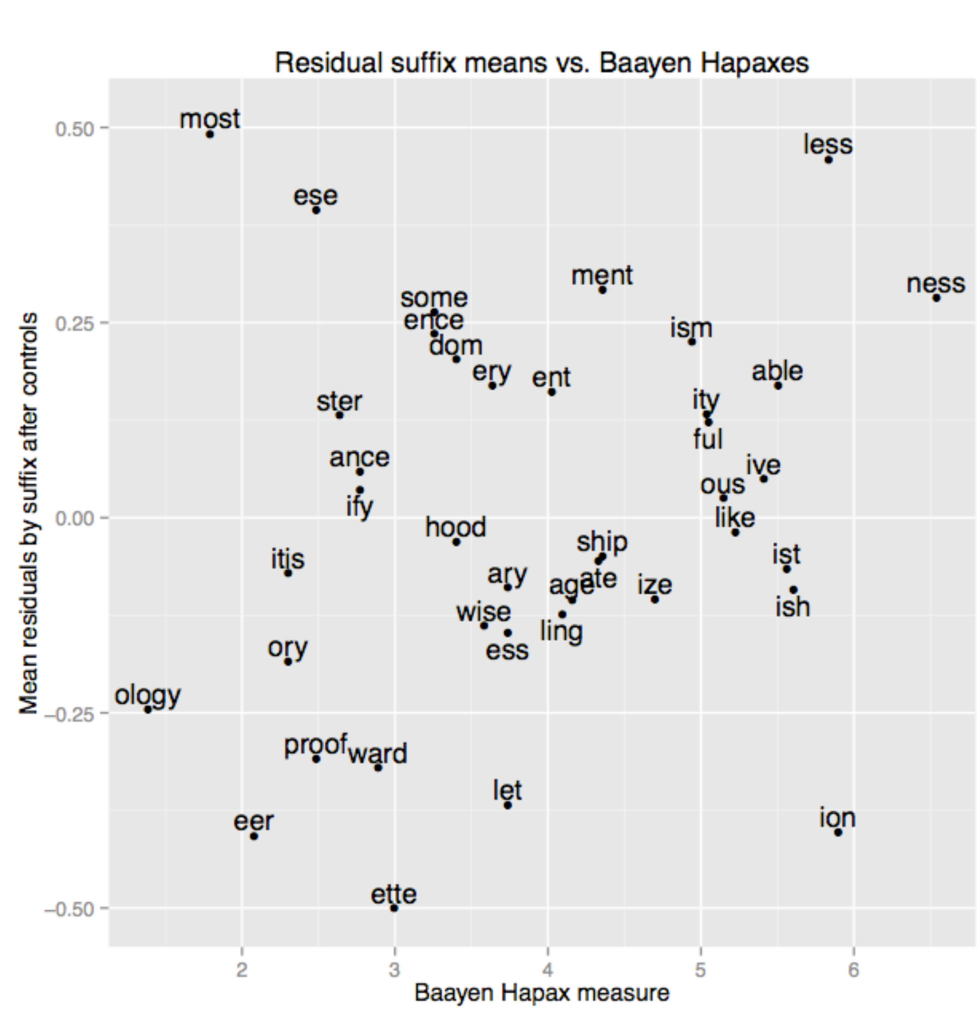
-To compare the productivity of various affixes, we use two theoretical measures of productivity: First, quantities computed from the Bayesian model of productivity known as Fragment Grammars (FG; O'Donnell 2011) and, second, Baayen's P*, A Good-Turing based based estimator (Baayen 1994).



Task

-Using Amazon's Mechanical Turk, we presented 206 participants with 105 word frames. Participants were asked to guess how many times a given pattern would occur in a 100,000 word novel. The 105 word frames presented to participants fell into one of four categories: (i) full-suffix frames like _ _ _ _ n e s s, (ii) partial-suffix frames like _ _ _ n _ s _, (iii) frames based on mono-morphemic words like r _ _ d (road, reed, etc.), and (iv) impossible frames like z s q _ _ _ .
-We ultimately analyzed 40 suffixes of 3 or more letters drawn from the corpus of O'Donnell (2011). For suffix-derived frames, each participant saw at most one instance of a given suffix. Partial-suffix frames were created by randomly deleting letters from full suffix frames. Mono-morphemic frames were sampled to have a wide spread of frames, from those with many possible completions like s _ _ _ _ to those with few, like b r i c _.

Results



Controls

- -log token frequency of the pattern (from Subtlex)
- -log type frequency of the pattern (from Subtlex)
- -number of missing letters in pattern
- -number of present letters in pattern
- -interaction between number of present letters and number of missing letters

Overall effect

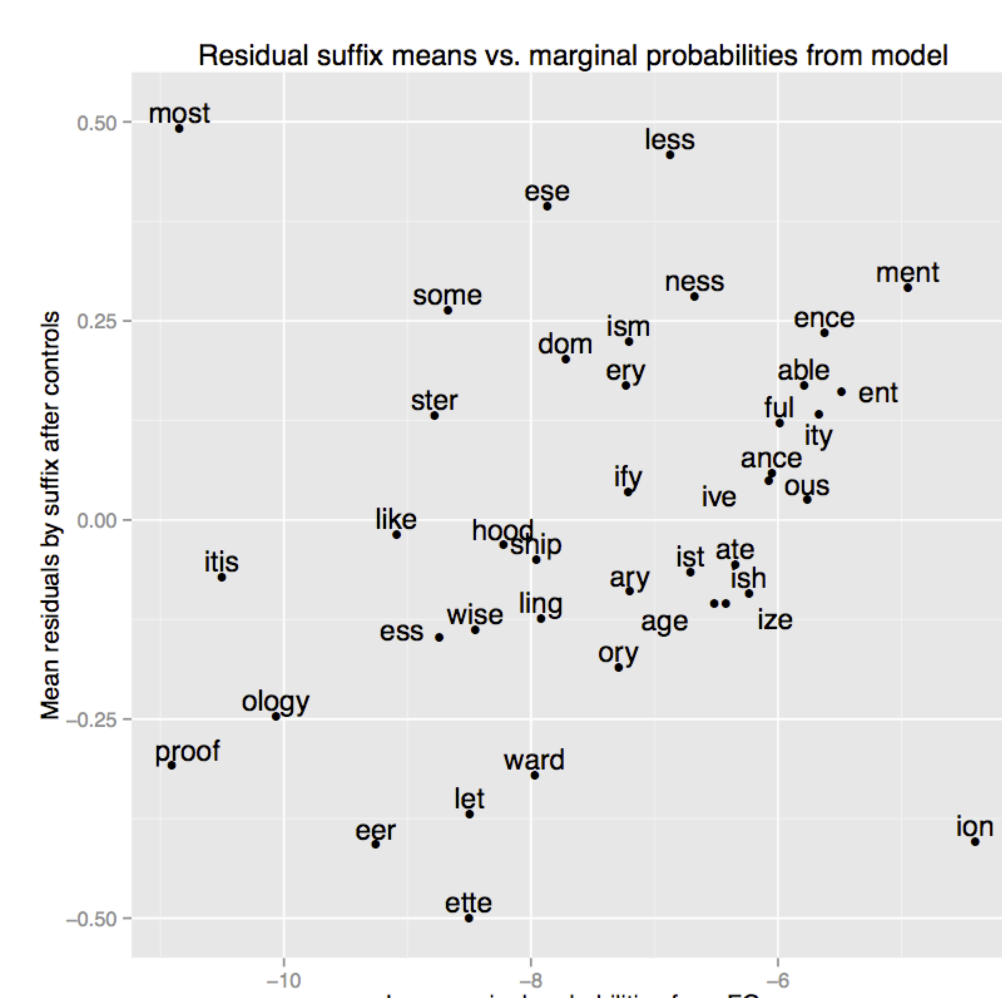
-We fit a mixed effect model with the above controls and a maximal random effect structure for subject. As predicted by T&K full suffix frames were inflated relative to a baseline for partial suffix frames while controlling for frequency (β=.97, t =20.43, chi2= 61, p< .001), and both full and partial frames were overestimated relative to mono-morphemic frames (β= .35, t =8.41, chi2 = 229, p < .001).

Productivity effect

-The FG model produces marginal probability scores that can be thought of as the probability of seeing a given suffix in either an existing word ***or*** in a novel word. Thus, the score can be thought of as a rational integration of the expected frequency of the suffix in the existing lexicon along with its expected probability of generalization.

- A higher marginal probability as estimated by FG was predictive of the residuals after regressing out the nuisance variables (β=.03, t=10.06, p < .0001).

-Baayen's P* was also significantly predictive of the residuals (β=.03, t=10.89, p < .0001).



Conclusion

-The role of morphological productivity in these estimates suggests that T&K's effect is likely caused by a "hallucinatory" effect of productive morphemes.
-Productive morphemes can give rise to unbounded numbers of novel forms, the presence of a productive suffix in isolation causes overestimation of the frequency of the pattern, with more productive morphemes leading to greater rates of overestimation.
-Thus, while T&K's result indicates that special purpose mechanisms can lead people to fallacious probabilistic reasoning, these mechanisms may be optimal in their intended domain, in this case, morphology.

Citations

Baayen, R. H. (2006). Corpus linguistics in morphology: Morphological productivity. In Ludeling, A. and Kyto M., editors, Corpus linguistics: An International Handbook. Mouton de Gruyter.

Bauer, L. (2001). Morphological Productivity. Cambridge U-boatniversity Press, Cambridge.

O'Donnell, T.J. (2011). *Productivity and Re-use in Language*. Harvard University. PhD dissertation.

Tversky, A., & Kahneman, D. (1973, September). Availability: A heuristic for judging frequency and probability. Cognitive Psychology, 5(2), 207–232.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. Psychological review, 90(4), 293.

Yang, C. (2010). Three factors in language variation. Lingua, 120:1160–1177.

Acknowledgments

We thank the CUNY 2013 reviewers, Leon Bergen, and the members of Tedlab.

Kyle's contact info:
kylemaho@mit.edu