Quantifying Perceived Morphological Relatedness

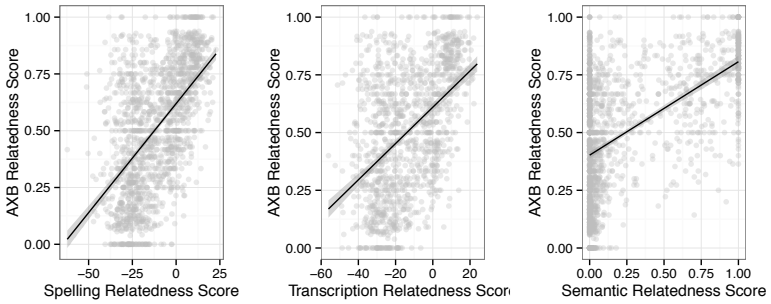
Kathleen Currie Hall, Claire Allen, Tess Fairburn, Kevin McMullin, Michael Fry, Masaki Noguchi

A key tool in a phonologist's toolkit is the use of morphophonological alternations, i.e., cases where a single morpheme has multiple phonological representations, varying by context. Most predictable phonological processes are known through their application in alternations. Whether a single morpheme is undergoing alternation is in many cases fairly trivial; e.g., inflections of a single verb provide easily interpretable data. But there are other cases in which determining whether two words are morphologically related is not so easy: does the morpheme *face* occur in *face-lift*? *facial*? *surface*? Etymologically speaking, the answer may be yes, but this does not necessarily reflect the morphological awareness of speakers. And yet, the recognition of such relatedness is crucial to the identification of alternations and thus to phonological processes.

Much of the morphological literature has focused on processing models (e.g., Stockall & Marantz 2006) and/or mostly dealt with inflection (e.g., Ackerman et al. 2009). In this paper, we present a quantification that does not attempt to isolate individual morphemes but rather gives an overview of perceived word similarity. An AXB discrimination task was used to elicit similarity judgments from English speakers. 180 individual key words (base morphemes, e.g, *press*) were matched with each of nine different words of the following types (given with examples): (1) inflection (*pressed*), (2) relatively transparent derivation (*pressure*), (3) relatively opaque derivation (*expressway*), (4) primary semantic sense (*push*), (5) secondary semantic sense (*media*), (6) rhyme (*mess*), (7) cohort (*preppy*), (8-9) unrelated words (*sofa*, *table*). For any given key word, two of these nine words were presented in orthographically; participants indicated which was "more similar" to the key. No instructions on what defines similarity were given. All combinations of all pairs were presented, randomly, with a single participant seeing only two samples of any given key word in 360 total trials. Results here are based on 26 pilot participants.

For each type of comparison word, the percentage of the time that that type was chosen when it was an option is given in the table below. Morphologically related forms are indeed interpreted as being most similar to the key, followed by semantically and then phonetically related forms. The behavioural (AXB) results were also correlated with similarity scores for spelling and pronunciation (cf. Khorsi 2012) and semantic relatedness (cf. Han et al. 2013), as shown below; R^2 values for the three are 0.31, 0.22, 0.30, respectively. Given, however, that morphological relatedness should really reflect a *combination* of sound and meaning relatedness, a logistic mixed-effects regression model of the behavioural results was developed. The model has a classification accuracy of 75% and a McFadden's R^2 value of 0.49. It crucially relies on spelling, pronunciation, and semantics as factors, suggesting that all three are relevant for predicting perceived morphological relatedness. This model can thus be used to independently approximate the degree to which any two words are morphologically related in the minds of speakers, which in turn can be used to determine the extent to which phonological processes that depend on alternations between morphologically related words are in fact psychologically viable.

Relation	*Percent*
Inflected	0.83
Derived1	0.74
Derived2	0.65
Meaning1	0.60
Meaning2	0.51
Rhyme	0.48
Cohort	0.38
Unrelated	0.17



References:

Ackerman, F., Blevins, J. P., & Malouf, R. (2009). Parts and wholes: Patterns of relatedness in complex morphological systems and why they matter. In J. P. Blevins & J. Blevins (Eds.), *Analogy in grammar: Form and acquisition* (pp. 54-82). Oxford: Oxford University Press.

Han, L., Kashyap, A., Finin, T., Mayfield, J., & Weese, J. (2013). UMBC_EBIQUITY-CORE: Semantic Textual Similarity Systems *Proceedings of the Second Joint Conference on Lexical and Computational Semantics*.

Khorsi, A. (2012). On Morphological Relatedness. *Natural Language Engineering*, 1-19.

Stockall, L., & Marantz, A. (2006). A single route, full decomposition model of morphological complexity. *The Mental Lexicon, 1*(1), 85-123.