WORD FORM SHAPES ARE CULTURALLY SELECTED FOR INDICATING THEIR MORPHOLOGICAL STRUCTURE

Andreas Baumann, Christina Prömer, and Nikolaus Ritt*

*Corresponding Author: nikolaus.ritt@univie.ac.at
Department of English and American Studies, University of Vienna, Austria

Listeners recognize morphologically complex word forms by their phonotactic shapes. For example, the final consonant clusters /md/ in *seemed*, /ld/ in *filled*, or /ks/ in *socks* function as clues that prompt decomposition. At the same time, experimental work (Post, Marslen-Wilson & Tyler 2008, cf. also Marslen-Wilson & Tyler 1997 & 1998) has shown that such recognition strategies are over-applied to word forms that are not actually complex, but simply happen to be shaped like complex ones. Thus, listeners attempt to decompose not only actual past tense forms such as *fill+ed* or *sign+ed*, or actual *plurals* such as *sock+s* but also simple words such as *build* or *find*, or *axe*, which delays processing, and leads listeners up garden paths. Obviously, however, such problems arise only when morphologically produced sound sequences have homophones among morphologically simple items. As we have seen, this is true of final clusters such as /nd/, /ld/ or /ks/. It is not true of clusters such as final /md/, /vd/, or /gz/, however. The latter occur only in past forms such as *seemed*, or *loved*, or in plurals, genitives, or 3sg forms such as *eggs*, *Meg's* or *digs*. Therefore, they signal complexity unambiguously and reliably.

If speakers are sensitive to the problems resulting from ambiguities between morphologically produced clusters (as in *fill+ed*) and lexically simple ones (as in *build*), they should be biased against the use of words that are ambiguous in that respect (Dressler & Dziubalska-Kołaczyk 2006; Dressler, Dziubalska-Kołaczyk & Pestal 2010). Since morphotactically ambiguous sequences are abundant in natural languages, however, the processing difficulties they produce do not seem to prevent successful communication and can only be slight. They nevertheless ought to be detectable in long-term language change, which results from vast numbers of communicative interactions and iterated learning events, known to

amplify even weak cognitive biases in language learning or language use (cf. Smith et al. 2017)

We have tested this hypothesis in a diachronic corpus study of highly ambiguous final clusters that arose through system-wide vowel deletion in unstressed syllables at the beginning of the Middle English period. In terms of data, it is based on the ECCE-database (ecce.univie.ac.at), which contains about 370.000 tokens of word forms ending in consonant clusters and covers the period from the middle of the 12th to the middle of the 18th century. The data are derived from the Penn Helsinki Parsed Corpora of Middle and Early Modern English (Kroch, Santorini & Delfs 2004, Kroch & Taylor 2000). In a statistical investigation of both type and token frequencies, we look for changes in the proportion of word forms that display either lexical or morphologically produced instances of final [sonorant]/d/, [sonorant]/t/, sonorant/z/, and sonorant/s/ clusters (as in *fill+ed* vs. *build*, *deal+t*,vs. *melt*, *name*+s vs. *Thames*, *mann+ys* vs. *prince*).

Our results show a significant trend towards cluster disambiguation: clusters in voiceless /t/ and /s/ have evolved to be increasingly indicative of morphological simplicity, while clusters ending in voiced /d/ and /z/ have come to signal complexity more reliably. The box chart in fig. 1 shows the difference for final sonorant+/t/ vs. sonorant-/d/ clusters on the level of types. The left box represents the distribution expected after schwa loss first produced final /nd/-clusters in words like *sinned* (/sind/ < /sinəd/) or *spilt* (/spilt/ < /spiləd/). The right one shows the distribution actually attested after the change.
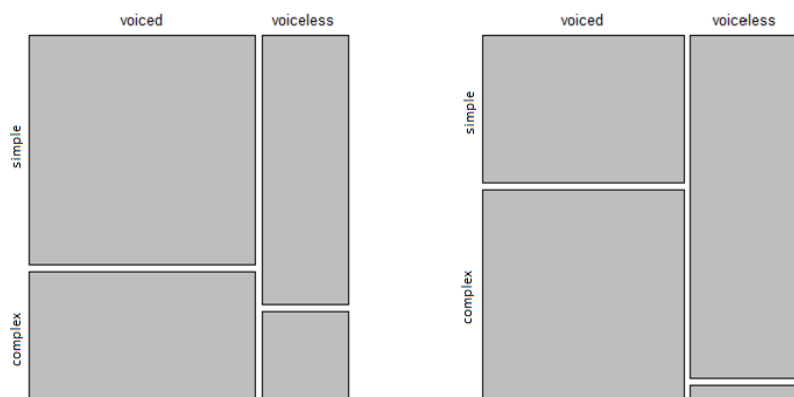

Figure 1. Changes in the distribution of SON+/t/ and SON+/d/ across simple and complex forms.

As fig. 2 below shows difference between sonorant+/t/ and sonorant+/d/clusters came to correlate significantly more strongly with the difference between simple and complex word forms.
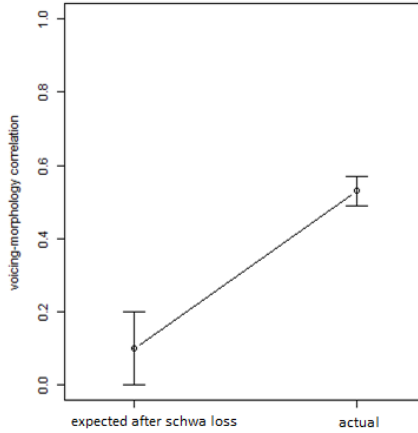

Figure 2. The evolving correlation between SON+/t/ vs. SON+/d/ and simple vs. complex words

One of the most interesting aspects of our results is that the disambiguation appears to be the cumulative effect of a set of rather diverse changes. They include sporadic sound changes (such as voicing of inflectional –s), morphological changes (the replacement of –V*nd* participles by gerund forms ending in –*ing*), as well as the adoption and loss of specific lexical items. We interpret this to imply a selection bias for word forms that signal their morphological structure reliably. Such a bias may represent a plausible – and so far unrecognized -macro-condition on the cultural evolution of natural languages that constrains the actuation and the spread of changes on various levels of linguistic organization and gives direction to the long-term development of natural languages.

References

Dressler, Wolfgang U. & Dziubalska-Kołaczyk, K. 2006. Proposing Morphono-tactics. *Wiener Linguistische Gazette* 73. 69-87.

Dressler, W. U., Dziubalska-Kołaczyk, K., & Pestal, L. 2010. Change and variation in morphonotactics. *Folia Linguistica Historica* 31. 51-68.

Kroch, A., Santorini B., & Delfs L. 2004. Penn-Helsinki Parsed Corpus of Early Modern English. http://www.ling.upenn.edu/hist-corpora/.

Kroch, A. & Taylor, A. 2000. Penn-Helsinki Parsed Corpus of Middle English. http://www.ling.upenn.edu/hist-corpora/.

Marslen-Wilson, W. D., & Tyler, L. K. 1997. Dissociating types of mental computation. *Nature* 387. 592–594

Marslen-Wilson, W. D., & Tyler, L. K. 1998. Rules, representations, and the English past tense. *Trends in Cognitive Science* 2. 428–435

Post, B., Marslen-Wilson, W. D., & Tyler, L. K. 2008. The processing of English regular inflections: Phonological cues to morphological structure. *Cognition* 109. 1–17.

Smith, K., Perfors, A., Fehér, O., Samara, A., Swoboda, K., & Wonnacott, E. 2017. Language learning, language use, and the evolution of linguistic variation. *Philosophical Transactions of the Royal Society B* 372 ,20160051.