Word frequency and 'fake cyclicity': a new analysis of English word stress preservation

Sarah Collie, University of Edinburgh, sejcollie@hotmail.com

Weak stress preservation can be seen at the left edge of English complex words with an odd number of pre-tonic syllables, e.g. *oríginal* → *orìginálity* (SPE; Liberman & Prince, 1977; Halle & Vergnaud, 1987; Benua, 1997; Pater, 2000; Marvin, 2003); *sensátional* → *sensàtionálity* (Kiparsky, 1979). Previous analyses of this phenomenon retain the core principle of SPE's cyclic stress application. In this paper, I show that cyclic stress application is incorrect.

Since SPE, weak stress preservation has been seen as a showpiece cyclic phenomenon: moving outwards, every morphological constituent constitutes a phonological domain so that secondary stress misapplies, e.g. *antícipate* → *antìcipátion*, not **àntïcipátion* as in monomorphemic *Lùxipalílla*. Recent analyses still respect this cyclic principle, e.g. Benua (1997) (in Optimality Theory) and Marvin (2003) (in Distributed Morphology). However, cyclicity as defined here is not always respected: weak stress preservation sometimes fails even though it is predicted by a word's morphological structure, e.g. *antícipate* → *àntïcipátion*. Under a cyclic analysis, these instances of preservation failure can only be accounted for by stipulating morphological structure: e.g. addition of a fused suffix to a bound root [[anticip]$_R$ ation]$_N$, rather than recognising each suffix as in [[[anticip]$_R$ ate]$_V$ ion]$_N$ (Ricardo Bermúdez-Otero, p.c.).

In this paper, I show that whether stress preservation is successful or not is significantly predicted by word frequency: instances of preservation failure like *antícipate* → *àntïcipátion* require no stipulation to account for them. I report on a new investigation into stress preservation behaviour for a controlled subset of words from Jones (2003) where second-syllable preservation is expected. Token frequencies for both embedding words (e.g. *anticipation*) and embedded words (e.g. *anticipate*) were collected from the CELEX database (Baayen et al., 1995). Statistical analyses indicated that stress preservation is more likely to fail if:

(i) The embedding word has a higher absolute frequency
(ii) The embedding word is more frequent than the embedded word, regardless of absolute frequency

Both results are anticipated in light of psycholinguistic research: more frequent words are less likely to retain properties (here, stress contours) of their embedded morphemes (Bybee, 1985); Hay (2001, 2003) shows that this weakening of morphological relationships becomes more likely as the embedding word's frequency increases relative to the embedded word's frequency.

My results can be handled by Bermúdez-Otero's (in preparation) notion of 'fake cyclicity': the pattern which appears to be the result of cyclic stress application is in fact the result of blocking among stored lexical entries. Lexical entries are argued to include a word's stress pattern. When an embedded word like *accélerate* is sufficiently frequent, its lexical entry is activated upon lexical access of the embedding word *acceleration*, and the default stress pattern found in monomorphemes is blocked to give *accèlerátion*, not **àccelerátion*. However, when the embedded word is not sufficiently frequent, e.g. *coóperate*, the default, monomorphemic stress pattern is not blocked and stress preservation may fail: *còoperátion*, not preserving *coòperátion*. This fake cyclicity analysis is compatible with dual-route race models of lexical access (e.g. Baayen, 1992): words may either be accessed via a decomposed route, reinforcing the preserving stress pattern, or directly via a whole-word route, favouring preservation failure.

My results cannot be understood in the traditional cyclic analysis: the cycle, unlike fake cyclicity, does not have the potential to be a probabilistic mechanism. In light of this outcome, the status of the cycle in the wider phonological context must be reappraised.