Commentary on Daelemans, Gillis, and Durieux

Jonathan Kaye[1]
(School of Oriental and African Studies)

The authors turn an Instance-Based Learning model (IBL) loose on some Dutch stress data. They come to a number of conclusions regarding the models that could underlie human stress acquisition, to wit, "... tree building operations proposed in learning theories for metrical phonology are not necessary for learning stress assignment" (p. 449) One assumes that this conclusion carries over to grid building as well. A second conclusion is that "... phonemic representations yield significantly better results [in learning Dutch irregular stress/JK] than the encoding in terms of syllable weights" (p. 449) In my opinion, the authors have failed to justify either conclusion.

Their work involves modeling the acquisition of Dutch stress. Specifically, the learning task is to assign Dutch input words to one of three categories: PEN, FIN, or ANT. Since systems very different from the Dutch one exist, the requirements of learning systems based exclusively on it can be carried over to the more general problem of human stress acquisition. Crucially, the authors do not consider the assignment of secondary stress or any of the other types of stress systems attested in the literature.

The authors consider three methods of encoding input forms: (1) syllable weights, (2) the "phonemes" of the rhyme projection, (3) a plain "phonemic" transcription of the word. The authors test these three methods with respect to Dutch stress placement. Research on human stress systems has failed to turn up examples where the segmental content of a form is relevant to general patterns of stress assignment. Statements of the form, "assign stress to the final nucleus unless it contains "i" in which case assign stress to the penultimate nucleus" are unattested. Accordingly, the conclusion of Daelemans et al. that "... the phonemic representations yield significantly better results than the encoding in terms of syllable weights" (p. 449) is surprising. Daelemans et al. note that for *regular* stress assignment (i.e. stress assignment based on syllable weight) the first encoding is the most successful. Encodings 2 and 3 only come into their own with respect to *lexical exceptions*. The authors fail to distinguish two very different operations: (1) the assignment of stress based on phonological structure and (2) the identification of lexical exceptions. While segmental content is irrelevant for the first task, it is essential for the second. It is clear that segmental material is required for the assignment of arbitrary marks ([−ex], LF, etc.) because we must know to which lexical items these diacritical marks are assigned. What emerges from the discussion is that the irrelevance of segmental material to stress assignment is supported by the current study. Indeed, there are many examples of stress systems where 100% of the lexical items belong to the "R" (regular) class. I know of no stress systems that are entirely based on patterns requiring "phonemic" information.

In conclusion, the restricted nature of the model deprives the theoretical claims of much of their relevance. Failure to distinguish stress assignment based on constituent structure vs. stress assignment as a property of certain lexical items attenuates the authors' claims about the relevance of the encoding methods. I believe we can remain secure in the belief that segmental material is irrelevant to stress assignment.

1 Department of Linguistics, School of Oriental and African Studies, Thornhaugh Street, Russell Square, London WC1H 0XG, U.K.