

## The role of prosody in constituent ordering

Does prosody influence constituent ordering? Our quantitative study of the English Dative Alternation ((1)) based on the Switchboard Corpus (spoken, 2,349 sentences, 37 distinct verbs) and a blog corpus ([www.blogspot.com](http://www.blogspot.com), written, 1,580 sentences, 16 distinct verbs) suggests that prosody matters. We present an OT model that uses prosodic constraints to predict the possible constituent orderings and their relative degrees of well-formedness.

We consider three effects. The first involves unstressed pronouns (e.g. *it*): unstressed pronoun themes usually do not occur in double object constructions ((2)); in dialects where they do, they preferably occur after pronoun goals ((3)) (Erteschik-Shir 1979, Hawkins 1995); unstressed pronoun goals are found in double object constructions with *donate*-verbs where other NPs are usually banned ((4)) (Bresnan and Nikitina 2003); unstressed pronoun goals are favored in double object constructions even with *give*-verbs ((5)); disfluencies like *uh, you know* are systematically absent before unstressed pronouns ((6)). This suggests two constraints: (i) unstressed pronouns cannot form a prosodic phrase alone ((2, 6)); (ii) unstressed pronouns are favored in (V NP) because they avoid stress clash with the verb ((3, 4, 5)). The second effect is “end weight”: phrases preferably occur in order of increasing length (e.g. Wasow 2002). We derive this from two constraints: (i) phrasal stress falls on the rightmost constituent (Chomsky and Halle 1968); (ii) word stresses preferably fall under phrasal stress. As a result, the number of word stresses in the constituent under phrasal stress is maximized. This predicts that unstressed function words (e.g. *a/an, of*) should not count for weight, a prediction supported by Switchboard where the number of lexically stressed words is a better predictor of phrasal weight than the number of words. This also predicts that languages with phrasal stress on the left should exhibit leftward weight effects (e.g. Chang and Yamashita 2001 on Japanese). The third effect involves verb length: verbs longer than one foot, e.g. *(do)(nate)*, typically do not alternate (Grimshaw 2005). This reflects two general constraints: (i) the verb forms a prosodic phrase with its adjacent complement (Inkelas and Zec 1995); (ii) prosodic constituents are preferably binary. Hence ((*give*) (NP)) NP is superior to \*((*do*)(*nate*) (NP)) NP because the latter contains a ternary prosodic phrase.

In our OT model, inputs are unordered; outputs are ordered and prosodified. Using OT software (Hayes, Tesar and Zuraw 2003, Anttila and Andrus 2007) we derive a large set of implicational predictions that are supported by quantitative evidence from both corpora. These predictions take the following form: if an input *i* permits a particular ordering and/or prosodification, so does input *j*. For example, *I gave my sister a book* (lexically stressed goal, double object) entails *I gave her a book* (lexically unstressed goal, double object). One of the model’s predictions is that two-foot verbs should exhibit more Heavy NP Shift than one-foot verbs. This is confirmed by the blog data where HNPS exclusively occurs with two-foot verbs, e.g. *reveal to her the truth* vs. \**give to her the book*.

Notation: [ ] = syntactic constituent; ( ) = prosodic constituent

- (1) (a) Celebrity status gave [Schwarzenegger] [options]. (Double Object)  
(b) Man gave [names] [to all the animals]. (Prepositional)  
(c) It is rather an abuse to administer [to laymen] [both forms]. (Heavy NP Shift)
- (2) (a) (Pat) (gáve it) (to Chris) (b) \*(Pat) (gáve Chrís) (it).
- (3) (a) ?I (gáve her) (it) (b) \*I (gáve my síster) (it)
- (4) (a) \*John (dónàted the chárity) (money).  
(b) ?They can get the gullible ones to (dónàte them) (money).
- (5) (a) *give-verb + pronominal goal + non-pronominal theme* →  
94.3% double objects in the blog corpus, e.g. *I gáve him the bók*  
(b) *give-verb + non-pronominal goal + non-pronominal theme* →  
26.6% double objects in the blog corpus, e.g. *I gáve Jím the bók*
- (6) (a) I'll (give you), you know, (two to three years).  
(b) \*I'll (give), you know, (you) (two to three years).