Phonological constraints on constituent ordering

Does phonology influence the ordering of meaningful elements (morphemes, words, phrases)? In general, the answer seems to be no (Pullum and Zwicky 1988), but an investigation into the quantitative distribution of constituents tells a different story. We report a quantitative study of the English Dative Alternation ((1)) in a prosodically annotated corpus based on data from www.blogspot.com (1,580 sentences, 16 distinct verbs). We conclude that prosody plays an active role in constituent linearization in English. The prosodic effects are mostly gradient and variable, yet entirely systematic.

Several observations suggest that prosody plays a role in the Dative Alternation. Here we summarize the main generalizations based on earlier work and our corpus study. First, lexically unstressed pronouns behave differently from other NPs: pronoun themes usually do not occur in double object constructions ((2)); in dialects where they do, they prefer to occur after pronoun goals ((3)) (Erteschik-Shir 1979, Hawkins 1995); unstressed pronoun goals are common in double object constructions with verbs like *lower*, *mutter*, *donate*, and *return* where other NPs are usually banned ((4)) (Bresnan and Nikitina 2003); and they are more frequent than other NPs in double object constructions even with verbs like *give* ((5)). These generalizations reflect two prosodic constraints: (i) unstressed pronouns cannot form a prosodic phrase alone ((1)); (ii) unstressed pronouns are preferred in the prosodic phrase (V NP) because they avoid stress clash with the preceding verb ((2, 3, 4, 5)). Second, heavy NPs tend to come last in the prepositional construction ((6)) (Wasow 1997, 2002). We derive this from two prosodic constraints: (i) phrasal stress falls on the rightmost constituent (= the Nuclear Stress Rule, Chomsky and Halle 1968); (ii) word stress preferably coincides with phrasal stress. The combined effect of these constraints is to maximize the number of lexical stresses in the constituent under nuclear stress, hence Heavy NP Shift (HNPS). This predicts that only the relative weight of the phrases should matter (Wasow 2002, cf. Jäger and Rosenbach 2004); that languages with sentence stress at the left edge should exhibit HNPS to the left (see McCawley 1977, Cinque 1993, Chang and Yamashita 2001 for Japanese); and that unstressed function words (e.g. *a/an*, *the*, *of*) should not count for weight. Third, the number of feet in the verb matters ((7)) (Fraser 1998, Grimshaw 2005). The core class of alternating verbs have one foot, e.g. (*give*), *a*(*ssign*); the non-alternating verbs have two or more feet, e.g. (*ex*)*(plain*) and (*do*)(*nate*). This is as expected if we assume that a verb forms a prosodic phrase with the adjacent complement (Inkelas and Zec 1995) and prosodic constituents are preferably binary. Thus, ((*give*) (NP)) NP emerges as prosodically superior to *((*do*)(*nate*) (NP)) NP because the latter contains a ternary prosodic phrase.

We present an optimality-theoretic model that predicts the possible constituent orderings for VPs of different prosodic types, as well as the relative well-formedness of each ordering. We consider 12 types of VPs with different prosodic properties; the output candidates consist of 8 possible realizations (= 4 linear orderings × 2 prosodic phrasings). Using optimality-theoretic software (Hayes, Tesar and Zuraw 2003, Anttila and Andrus 2006) we derive a large set of implicational predictions that are independent of constraint rankings and that emerge quantitatively in the observed corpus frequencies. These predictions take the following form: if an input *p* permits a double object / prepositional / HNPS construction, so does input *q*, e.g. *I gave my sister a book* (lexically stressed goal, double object) → *I gave her a book* (lexically unstressed goal, double object). One novel prediction is that two-foot verbs should exhibit more HNPS than one-foot verbs. This is borne out by the corpus 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] (Blogspot)
(b) Man gave [names] [to all the animals]. (Blogspot)

(2) (a) (Pat) (gave it) (to Chris)
(b) *(Pat) (gave Chris) (it).

(3) (a) ?I (gave her) (it)
(b) *I (gave my síster) (it)

(4) (a) *I lowered John the box.
Buddha lowered him the silver thread of a spider. (Bresnan and Nikitina 2003)
(b) *Susan muttered Rachel the news.
She muttered him a hurried apology. (Bresnan and Nikitina 2003)
(c) *John donated the charity money.
They can get the gullible ones to donate them money. (Google)
(d) *John returned the government the money.
Judas returned them the money (Google)

(5) (a) I (gave her) the book (94.3% of VPs with one-foot verb + pronoun + NP)
(b) I (gave my sister) the book (26.6% of VPs with one-foot verb + NP + NP)

(6) A staff sergeant is explaining [to the men] [the rules of the Geneva Convention].
(Blogspot)

(7) (a) They (gave) the church money.
(b) *They (do)(nated) the church money.