

A non-syntactic template for syntactic noun incorporation

Michael J. Houser and Maziar Toosarvandani

University of California, Berkeley

Baker 1996:25-30 observes that syntactic noun incorporation places the incorporated noun (IN) before the verb root, regardless of the language's basic word order. Baker's explanation for this ordering is that incorporation involves a lower morpheme raising to left-adjoin to a higher one. The same principle predicts that tense morphemes should follow verb roots while agreement affixes should precede them (or, more generally, tense and IN should be on different sides of the verb root, agreement and IN on the same side). Baker et al. 2005 retract the claim that preverbal ordering is a universal since some counterexamples are attested.

Despite the importance of ordering issues in typology, the position of IN's has not had a systematic cross-linguistic survey. We survey incorporation type (lexical/syntactic, productive/unproductive), IN position, and tense and agreement position in one language per family of those known to have incorporation and a geographically matched sample of languages without incorporation, and use another 200-language genealogically and geographically diverse sample for general morphosyntactic comparison.

IN's are indeed usually preverbal (24 vs. 8 languages), regardless of the language's word order and equally for both syntactic and lexical incorporation. The 8 languages with postverbal IN are nearly all VO; in these languages minimal noun/verb distinctions and isolating IN (i.e. with some of the phonological properties of a separate word) are much more common than in languages with preverbal IN. Unproductive incorporation is never postverbal. Isolating incorporation is rarely unproductive. Tellingly, in two languages there is unproductive preverbal incorporation and productive postverbal incorporation. Thus Baker's claim is statistically very strong: preverbal IN is universally preferred, although not categorical. However, since conformity is strongest in unproductive and fossilized contexts and counterexamples occur precisely where syntactic derivation should be most transparently reflected – in productive and/or isolating patterns – something other than movement in syntactic derivation must be responsible for the order.

Position of tense and agreement morphemes shows a modest preference for what Baker predicts, not strong enough to support his explanation.

Based on a survey still in progress, preverbal position for the nominal or nounlike element also appears to be preferred in lexicalized complex predicates with light verbs, in incorporated adpositions and/or PP's, and in nominalized compounds with an undergoer such as English *scrapper*, *housewarming* – again largely independent of clause word order.

We conclude that the output of syntactic incorporation is probabilistically attracted to a universally favored ordering principle that shapes noun-verb combinations of other kinds as well and is independent of syntactic ordering. The theoretical status of templates in typology and formal theory is not settled, but preverbal IN may be an example of one. In diachronic terms, older, unproductive incorporation universally follows the template, while newer and/or non-affixal incorporation sometimes follows the language's word order rather than the template. That is, the longer incorporation has been in the language the greater the chance that it will conform to the template. Paradoxically, while the template is head-final, head-initial morphosyntax appears to favor the appearance of incorporation in the first place.