



PROJECT MUSE®

Asymmetries in the prosodic phrasing of function words: Another look at the suffixing preference

Nikolaus P. Himmelmann

Language, Volume 90, Number 4, December 2014, pp. 927-960 (Article)

Published by Linguistic Society of America

DOI: 10.1353/lan.2014.0105

LANGUAGE
A JOURNAL OF THE LINGUISTIC
SOCIETY OF AMERICA

CONTENTS OF VOLUME 90	
Editorial	xi
Asymmetries in the prosodic phrasing of function words: Another look at the suffixing preference	927
Review Article	961
Book Reviews	961
Index	961

➔ For additional information about this article

<http://muse.jhu.edu/journals/lan/summary/v090/90.4.himmelmann.html>

ASYMMETRIES IN THE PROSODIC PHRASING OF FUNCTION WORDS:
ANOTHER LOOK AT THE SUFFIXING PREFERENCE

NIKOLAUS P. HIMMELMANN

Universität zu Köln

It is a well-known fact that across the world’s languages there is a fairly strong asymmetry in the affixation of grammatical material, in that suffixes considerably outnumber prefixes in typological databases. This article argues that prosody, specifically prosodic phrasing, plays an important part in bringing about this asymmetry. Prosodic word and phrase boundaries may occur after a clitic function word preceding its lexical host with sufficient frequency so as to impede the fusion required for affixhood. Conversely, prosodic boundaries rarely, if ever, occur between a lexical host and a clitic function word following it. Hence, prosody does not impede the fusion process between lexical hosts and postposed function words, which therefore become affixes more easily.

Evidence for the asymmetry in prosodic phrasing is provided from two sources: disfluencies, and ditropic cliticization, that is, the fact that grammatical proclitics may be phonological enclitics (i.e. phrased with a preceding host), but grammatical enclitics are never phonological proclitics. Earlier explanations for the suffixing preference have neglected prosody almost completely and thus also missed the related asymmetry in ditropic cliticization. More importantly, the evidence from prosodic phrasing suggests a new venue for explaining the suffixing preference. The asymmetry in prosodic phrasing, which, according to the hypothesis proposed here, is a major factor underlying the suffixing preference, has a natural basis in the mechanics of turn-taking as well as in the mechanics of speech production.*

Keywords: affixes, clitics, language processing, turn-taking, grammaticization, explanation in typology, Tagalog

1. THE SUFFIXING PREFERENCE. The fact that suffixing is the most common morphological technique in the languages of the world has been repeatedly noted in the literature since Sapir (1921:70) at the latest. Quantitative data in support of this fact have been presented in Greenberg 1957, Hawkins & Gilligan 1988, Bybee et al. 1990, Cysouw 2009, and Dryer 2011a–e. Table 1 shows the distribution of grammatical elements appearing on verbs, including tense, aspect, mood, and person markers and valence-changing markers (passive, applicative, etc.), from Bybee et al. 1990:5.

	PREPOSED	POSTPOSED	TOTAL
affixes	426	1,236	1,662
function words	386	316	702
TOTAL	812	1,552	2,364

TABLE 1. The suffixing preference in verbal grammatical elements from seventy-one languages (from the database of Bybee et al. 1990:5).

The first line in Table 1 provides clear evidence for the suffixing preference: suffixes are about three times more numerous than prefixes in this database. The rest of the table, which includes function words expressing verbal grammatical categories, sug-

* This article has been in the making for many years and preliminary versions have been presented on numerous occasions, including the MPI Nijmegen Formal Colloquium Series (2002), the 6th biennial meeting of the Association for Linguistic Typology in Padang (2005), and invited lectures at the Universities of Bielefeld, Bochum, Düsseldorf, Erfurt, and Münster. I am very grateful for the many comments received on these occasions, in particular for the thorough critiques by Ann Cutler and Matthew Dryer. The current written version has profited considerably from comments by Uta Reinöhl, Sonja Riesberg, Volker Unterladstetter, and two anonymous referees for *Language*. Very special thanks to associate editor Kie Zuraw for her detailed and concise instructions for revision. Special thanks also to Jan Strunk for help with the figures.

gests that the disproportionately high number of suffixes may be the result of two different and possibly unrelated tendencies. First, there is a strong overall tendency to postpose grammatical elements: the ratio of preposed to postposed grammatical elements in the Bybee et al. database is roughly one to two (812 to 1,552). This is referred to here as the *POSTPOSING TENDENCY*. Second, postposed grammatical elements are typically affixes, while preposed grammatical elements are equally likely to be affixes or function words. This is the *SUFFIXING PREFERENCE*, in the narrower sense.

This article is concerned only with the suffixing preference proper. The term ‘suffixing preference’ is used only in reference to it, except when referring to the earlier literature on this topic. As is seen below, in earlier work this term is often used to refer to the fact that suffixes greatly outnumber prefixes in the languages of the world and therefore refers to the combined outcome of these two separate (strictly speaking) tendencies. Distinguishing these tendencies is particularly important when attempting to explain the preference for suffixes in natural languages, because the explananda are not identical. That is to say, the task of explaining the tendency to postpose grammatical elements is not the same as the task of explaining why postposed grammatical elements are more likely to be affixes than preposed ones.

The postposing tendency is partially due to the fact that verb-final (OV) languages are heavily and consistently postposing, as illustrated in Table 2, where postposed grammatical elements in OV languages outnumber preposed ones by roughly four to one (1,018 to 233). In verb-initial and -medial (VO) languages, by contrast, the two positional types are almost equally frequent (579 preposed to 534 postposed elements). But note that VO languages also contribute to the postposing tendency because they are not heavily preposing, and hence do not counterbalance the strong postposing tendency found in OV languages.

	OV		VO		TOTAL
	PREPOSED	POSTPOSED	PREPOSED	POSTPOSED	
affixes	158	813	268	423	1,662
function words	75	205	311	111	702
TOTAL	233	1,018	579	534	2,364

TABLE 2. The distribution of pre- and postposed grammatical elements in OV versus VO languages (from Bybee et al. 1990:6).

As for the suffixing preference proper, there are no major differences between OV and VO languages. In both language types, 80% of the postposed grammatical elements are affixed (see columns 3 and 5 in Table 2). There appears to be a difference in the affixation rate of preposed elements, in that more than 60% of the preposed grammatical elements are affixes in OV languages, but less than 50% in VO languages (see columns 2 and 4 in Table 2). As Bybee and colleagues (1990:7f., 13–18) point out, this difference is due to the fact that V-medial (SVO) languages appear to be particularly reluctant to affix preposed grammatical elements. Only 40% of the preposed grammatical elements in V-medial languages are affixes in their sample. We return to this issue in §5 below.

One point that has hardly been addressed in the literature on this topic is the fact that the suffixing preference is not attested to the same degree across ALL grammatical categories. On the one hand, as further discussed and exemplified in Cysouw 2009 (to my knowledge, the only published work to date that addresses this issue), the suffixing preference is particularly strong in the case of nominal plural, tense-aspect, and case marking, with suffixes accounting for over 80% of the attested affixes in these categories. Table 3 illustrates this point with data from Dryer (2011b,c,e). Note that this

table provides figures for LANGUAGES that primarily or exclusively make use of prefixes or suffixes for the category in question, while the two tables above (based on Bybee et al. 1990) give figures for grammatical elements. Hence the data from these two sources are not directly comparable.

	PREFIXING	SUFFIXING	TOTAL
	LANGUAGE	LANGUAGE	(LANGUAGES)
CASE MARKERS	38 (7.8%)	452 (92.2%)	490
TENSE-ASPECT	153 (18.6%)	668 (81.4%)	821
PLURAL	126 (19.7%)	513 (80.3%)	639

TABLE 3. Grammatical categories that show a strong suffixing preference (data from Dryer 2011b,c,e).

On the other hand, the preference for suffixes in the domain of person marking on verbs (pronominal subject and object affixes) is much less pronounced. In the Bybee et al. database, the ratio is 354 suffixes (59.6%) to 240 prefixes (40.4%) (Bybee et al. 1990:9, 13, 15).¹ Cysouw (2009) claims that in languages with particularly small person-marking paradigms (only two to four categories being distinguished), there is actually a PREFIXING preference in his data set.

Obviously, convincing explanations for the suffixing preference will also have to address the issue of why this preference is more clearly seen in some categories than in others, and more specifically, why it is rather weak, if it exists at all, in the domain of person marking. We return to this issue in §4.

While there is a well-established terminology for distinguishing preposed and postposed affixes and clitics (i.e. prefixes vs. suffixes and proclitics vs. enclitics), there are no well-established terms for making this distinction for (nonclitic) function words and for grammatical elements in general. This article follows Bybee et al. 1990 in using PREPOSED and POSTPOSED to fill this terminological gap, and thus speaks of PREPOSED vs. POSTPOSED FUNCTION WORDS and GRAMMATICAL ELEMENTS. The latter term is used as the superordinate term for affixes, clitics, and nonclitic function words (hence e.g. preposed grammatical elements comprise prefixes, proclitics, and preposed (nonclitic) function words).² This terminological provision is less than optimal as pre- and postposing may wrongly suggest movement of some sort. No movement of any kind, however, is intended by use of these terms. They simply denote the position of a grammatical element as either before or after its morphosyntactic host.

The next section briefly reviews the explanations for the suffixing preference proposed so far and develops the hypothesis that differences in the prosodic phrasing of preposed and postposed function words are of major import in this regard. Evidence for this hypothesis is provided in §3, and a scenario for explaining the suffixing preference based on the evidence provided by prosodic phrasing is sketched in §4.

2. EXPLAINING THE SUFFIXING PREFERENCE. The suffixing preference (broadly conceived) has become the primary example of a certain type of explanation in linguistics, that is, the attempt to motivate aspects of grammatical structure (and here specifically morphological structure) by direct recourse to findings from psycholinguistics. This line of research was inaugurated in a series of papers by Cutler, Hawkins, and Gilligan (Cut-

¹ The online edition of the *World atlas of language structures* (Dryer & Haspelmath 2011), from which Dryer's data in Table 3 are quoted, does not include figures for pronominal affixes on verbs.

² Bybee and colleagues (1990 and elsewhere) use GRAM as an abbreviation for grammatical elements.

ler, Hawkins, & Gilligan 1985, Hawkins & Cutler 1988, Hawkins & Gilligan 1988). Their explanation builds on two lines of psycholinguistic research. On the one hand, there is very strong evidence that word onsets (rather than endings or intermediary segments) are the most important access cues in word comprehension. This finding provides the basis for models of word recognition, such as the cohort model by Marslen-Wilson and colleagues, which heavily relies on left-to-right (earlier-to-later) processes (see, for example, Marslen-Wilson 1987 and many subsequent works). On the other hand, there is also good evidence that stems and affixes are processed separately, at least at some level (Hawkins & Cutler 1988:305). The central hypothesis explaining the suffixing preference, then, is the claim that 'speakers and listeners process stems before affixes' (Hawkins & Cutler 1988:306), which means that the earlier lexical stems become available the better.

The fact that many languages have prefixes and that there are even a few exclusively prefixing languages is explained by a principle that is in competition with the suffixing preference. This is the HEAD ORDERING PRINCIPLE, which states that affixes are ordered on the same side of their lexical bases as syntactic heads are relative to their complements (specifically, adpositions relative to their NP complement and verbs relative to their direct object argument). This principle thus predicts that prefixes should be the norm in VO and/or preposition-NP languages, while suffixes should be the norm in OV/NP-postposition languages (which is factually not correct, as seen in Table 2).

Hall (1992) provides a thorough critique of the proposal by Cutler, Hawkins, and Gilligan. His main concern is with the nature of explanations in linguistics, in particular functional explanations that invoke an extragrammatical feature or constraint in order to explain linguistic structures. Building on the classic paper by Clark and Malt (1984), he argues that most explanations of this kind fail to establish an explicit link between the explanatory feature and the structural feature to be explained. In order to provide such a link, it is necessary to show how the postulated constraint shapes linguistic structures in language use, that is, in concrete events of speaking, usually over extended periods of time. In the present case, the issue is how exactly the hypothesized preferred computational order in word recognition (stem before affixes) accounts for the fact that suffixes are more common than prefixes.

Any full explanation for the suffixing preference thus has to take into account the genesis of affixes. While it appears to be obvious that the preferred computational order disadvantages prefixed forms, it is far from clear how it actually impedes the development of such forms. Therefore, in order to link the postulated constraint (preferred computational order) to the structural feature it is supposed to explain (the suffixing preference), one has to identify factors that accelerate the affixation of postposed grammatical elements, or factors that impede the affixation of preposed grammatical elements or cause prefixed items to fall out of use in favor of other constructions.

It is well known that, historically, affixes typically develop from full independent lexemes.³ A standard example is the development of the French future suffixes from a Latin construction consisting of the matrix verb *habēre* 'have, own' plus infinitive: thus French (*je*) *chanter-ai* 'I will sing' historically derives from *cantāre habeo* 'I have to sing'. One parameter in this grammaticization process that is important for the present topic is the parameter of COALESCENCE, which refers to the degree to which the for-

³ The historical source and path of development in fact provides a possible explanation for the considerable number of 'exceptions' to Cutler and colleagues' explanation of the suffixing preference, as amply illustrated in Mithun 2003.

merly independent lexeme becomes phonologically dependent on its host. The cline in 1 shows the major stages in this process (see also Hopper & Traugott 1993:6f., Lehmann 1995:131–40).

(1) CLINE OF INCREASING COALESCENCE IN GRAMMATICIZATION PROCESSES:

lexeme → ‘heavy’ function word → clitic function word → affix → inflectional formative (→ 0)

Three stages in this cline are of major import for the current investigation. ‘HEAVY’ FUNCTION WORDS are items that phonologically behave like content words (nouns and verbs, for example), but express grammatical meanings. Examples include secondary prepositions (*despite, between, among*, etc.), conjunctions (*because, while*, etc.), and some (forms of) auxiliaries (*must, do/did/does*, etc.). CLITIC FUNCTION WORDS (or CLITICS, for short) also express grammatical meanings, but need a phonological host. The relation to this host is a purely phonological one, which means that clitics are generally not selective with regard to semantic or morphosyntactic properties of their hosts. While they may trigger or undergo phonological alternations in accordance with the phonological environment, they do not trigger allomorphies and generally also do not have multiple allomorphs. AFFIXES form not just a phonological but also a morphological unit with their hosts. The morphological dependency of affixes may be manifest in allomorphic alternations and the fact that the host may structurally require the presence of one or more affixes to function as a grammatical unit.

The distinction between these three different types of grammatical elements, in particular the one between clitics and affixes, is not always easy to apply in specific instances, as widely discussed in the literature (Zwicky & Pullum 1983, Spencer & Luís 2012:108–12, *passim*). Intermediary types and overlaps are in fact to be expected, if the development along the cline in 1 is gradual, as assumed in Hopper & Traugott 1993 and most other work on grammaticization processes. For current purposes, however, these problems of delimitation are not of direct import. The important point is that there are historically attested cases where all specialists agree that a given grammatical marker at some stage of its development was a clitic and at a later stage became an affix. The future markers in French and most other Romance languages are prime examples (Diez 1838:114f.).

Another aspect of importance for current concerns is the widely attested variability between ‘heavy’ and clitic function words. Well-known examples include English auxiliaries such as *will* → *’ll* or *have* → *’ve* or the negative *not* → *n’t*. The fact that grammatical elements that typically occur as clitics very often allow for alternative, nonclitic realizations is a major concern in §4 below. Note that this variability provides for a rarely noticed difference between clitics and affixes. Affixes never allow for alternative realizations as independent phonological words.

Although a change in position may occur in grammaticization processes (e.g. a preposed function word becomes a postposed one), it is generally the case that grammatical elements do not change position relative to their hosts once they have reached the clitic stage.⁴ Thus, proclitics usually become prefixes and enclitics usually become suffixes (see also Bybee et al. 1990:3, *passim*).⁵ Given this widely attested path of development, the explanandum with regard to the suffixing preference is this (see Hall 1992:96).

⁴ This, in principle, also holds for second-position clitics, where, however, morphosyntactic and phonological hosts typically differ, which usually impedes further development along the coalescence cline.

⁵ In fact, matters are more complicated with regard to proclitics. These may become enclitic on a host word they are not morphosyntactically related to. See §3.2 below for examples and discussion.

- (2) EXPLANANDUM: Why do postposed function words become affixes more often than preposed function words?

There are various factors that may influence the coalescence between grammaticizing function words and their hosts. One obvious factor is EXCLUSIVE ADJACENCY (Hall 1988:334, 1992:162). In order for function words to become affixes it is necessary that function words and their hosts regularly (and frequently) occur immediately adjacent to each other. This condition is not met when other elements can intervene between function word and host. In English and other Germanic languages, for example, articles, quantifiers, and adjectives regularly come between prepositions and nouns (as in example 3), which is the likely reason why there is absolutely no evidence for the development of case PREFIXES in these languages.⁶

- (3) [.75] and empties the pears ... **into** [.55] **big .. baskets**. (Chafe 1980:310, story 11)⁷

Hall (1988:334, 1992:162) notes that there is no evidence that preposed and postposed function words clearly differ with regard to exclusive adjacency. That is, in general, preposed function words do not appear to be positionally more variable than postposed ones. We briefly return to this issue in §5.

Hall (1992:Ch. 6) hypothesizes that the ‘reluctance’ of preposed function words to coalesce is due to the preferred computational order of processing stems before affixes, as in the Hawkins & Cutler 1988 approach. In his model, preposed function words are continually involved in a ‘flirting’ process with the stem, attempting to fuse with it. But unlike in the case of postposed function words, which are also involved in the same type of process, the ‘flirt’ never results in actual fusion, because the assumed additional processing cost involved in having to separate prefix from stem interrupts the coalescence process.

This explanation involves a rather complex set of hypotheses and very much depends on the correctness of the assumption that, universally, stems are processed before affixes in word recognition. One very basic problem of this assumption is that it implies that dealing with prefixes poses a constant processing handicap for speakers and raises the question of why prefixing should be attested at all in the languages of the world. Given that there are languages that are heavily or even exclusively prefixing (Na-Dene (Navaho), northwest Caucasian languages (Adyghe, Kabardian), languages of northern Australia, Mon-Khmer), one would like to be able to point to competing factors that could motivate languages to make use of prefixing as a morphological strategy despite the fact that they pose a processing handicap. The head ordering principle does not constitute such a factor, as convincingly argued by Hall (1992:Ch. 3). It is fraught with var-

⁶ Many of the examples in the present article are drawn from the work of Wallace Chafe and colleagues and use the transcription conventions given in Chafe 1980:301 and, more fully, in Du Bois et al. 1993. Updates and expansions can be found at <http://www.linguistics.ucsb.edu/projects/transcription/representing>. For current purposes, the most important conventions are the following:

- line intonation unit (IU)
- , continuative intonation at end of IU
- . terminative intonation at end of IU
- (1.2) timed pause (in seconds and tenths of seconds); in the Pear stories (Chafe 1980), square brackets [] are used for this purpose. Boundary pauses, occurring in between two IUs, are placed at the beginning of the line representing the IU after the pause.
- .. micro pause, too short to be measured
- truncated word
- truncated intonation unit

⁷ Boldfacing in all examples from the Pear stories of Chafe 1980 has been added by the author.

ious problematic assumptions, in particular the assumption that affixes are heads. Obviously, it would be more convincing to find an explanation based on factors that apply equally to all types of constructions (and languages), but have differing consequences with regard to preposed and postposed function words.

In this regard it is useful to note that exclusive adjacency is usually understood to be a purely positional requirement: a function word and host must regularly occur immediately adjacent to each other. It is clear, however, that mere positional adjacency is not enough. There also has to be prosodic integration for further coalescence to occur. Bybee and colleagues (1990:29) remark that 'in order for a gram to fuse with a stem it must occur contiguous to the stem with sufficient frequency and be in the same phonological phrase or intonation unit as the stem'. Essentially, this is already implied in the coalescence cline given in 1: in order to become an affix, a function word has to go through a clitic stage that is defined by its forming a phonological unit with its host.

A promising starting point for explaining the suffixing preference, then, is the hypothesis that preposed and postposed function words differ with regard to prosodic integration, as given in 4.

- (4) HYPOTHESIS: The grammaticization of a function word into an affix presupposes that the function word and its lexical host regularly form a prosodic unit (a prosodic word or phrase). This requirement is often not met for preposed function words, because prosodic word and phrase boundaries may occur between preposed function words and their lexical hosts with sufficient frequency so as to impede the fusion process. Postposed function words are much more rarely, if ever, separated from their lexical hosts by such boundaries.

It should be obvious that this hypothesis, if substantiated, is only a first step toward a comprehensive explanation of the suffixing preference. It simply shifts the explanandum to another level, that is, why preposed and postposed function words behave differently with regard to prosodic integration. We return to this issue in §4, but turn first to the task of providing evidence for the hypothesis in 4.

The hypothesis further develops and refines the explanation for the suffixing preference proposed by Bybee and colleagues (1990), who do not further investigate issues of prosodic integration, but make a related observation about clause boundaries. They claim that function words at clause boundaries, that is, preceding V in a V-initial language and following V in a V-final language, tend to become affixes, while clause-medial function words are more reluctant to fuse. The latter include not only pre- and postposed grammatical elements in V-medial languages, but also postposed markers in V-initial and preposed markers in V-final languages. Clause-medial function words only coalesce with their hosts if semantically relevant. As they put it: 'Our proposal for explaining the suffixing predominance, then, is that grams at clause boundaries tend to affix at a very high rate, while the rate of affixation for clause internal grams is determined by their meaning and relevance to the verb' (Bybee et al. 1990:34). This in turn means that the suffixing preference is primarily due to the fact that V-final languages vastly outnumber V-initial languages in their sample. Additionally, the V-final languages in their sample 'are more consistent' (*loc. cit.*) in positioning their function words at the right edge of the constituent they belong to, that is, clause-finally in the case of function words pertaining to the verb or the clause.

This account raises the question of why clause boundaries should be particularly conducive to the fusion of function words with their lexical hosts. It is argued here that it is precisely the fact that clause boundaries are generally also prosodic boundaries that is

relevant in this regard, because function words at clause boundaries will typically form lower-level prosodic constituents with their hosts.

3. FUNCTION WORDS AND PROSODIC BOUNDARIES. It is widely agreed that function words typically allow for a number of different prosodic realization options (see Anderson 2005 and Spencer & Luís 2012 for recent overviews). These include the option of being prosodized as prosodic words of their own, provided that they fulfill basic constraints such as minimal size (e.g. being disyllabic) or being associated with focus. Other options are different types of cliticization. Thus, for example, Selkirk 1995 and many subsequent works make a three-way distinction between free, internal, and affixal clitics, which are distinguished by different positions in a prosodic constituency tree. Regardless of its prosodic type (clitic or nonclitic), a function word typically forms a higher-level prosodic constituent with its lexical host (a prosodic word in the case of clitics, a phonological phrase in the case of function words prosodized independently as prosodic words).⁸ Two phenomena that deviate from this prototypical scenario are our concern here.

One phenomenon is the well-known possibility of a systematic mismatch between prosodic and syntactic structure along the following lines: a function word that morphosyntactically is a PROCLITIC behaves phonologically as an ENCLITIC, as, for example, when a preposition or preposed case marker is regularly phrased with the preceding content word rather than with its proper morphosyntactic host. Significantly for current purposes, the reverse does not seem to occur (i.e. a morphosyntactic ENCLITIC behaving phonologically as a PROCLITIC). This is further exemplified and discussed in §3.2 below.

The other phenomenon pertains to the possibility that in spontaneous speech, the prosodic unit of a (clitic) function word and its lexical host may be interrupted by a disfluency, for example, a hesitation pause. As is illustrated in §3.1 below, this is fairly common for preposed function words, but is not found for postposed function words.

On first sight, these two phenomena would appear to be of very different natures. Disfluencies properly belong to the domain of spontaneous speech and are extragrammatical inasmuch as they are not part of regularly formed structures. But this does not mean that they are completely haphazard and lack grammatical structure altogether. There are systematic aspects to their occurrence and their implementation that provide important evidence for processes in speech production, as amply documented in the psycholinguistics literature (e.g. Goldman Eisler 1968, Levelt 1989, Clark & Wasow 1998, Clark & Fox Tree 2002). As further discussed and illustrated in §§3.1 and 4, they require monitoring and planning on the part of the speaker and crucially involve the readjustment of prosodic word and phrase boundaries.

Mismatches between the prosodic and syntactic alignment of clitics, by contrast, properly belong to the realm of grammar. We are dealing here with a systematic property of regularly formed expressions. The two phenomena are investigated together here because both of them evince the same asymmetry with regard to preposed and postposed (clitic) function words. More specifically, it is suggested that they both are natural outcomes of the incremental left-to-right nature of speech production, as further detailed in §4.

The following two sections provide further evidence for the two phenomena and discuss their relevance for the suffixing preference.

⁸ The alternative of positing a clitic group as a level of its own (as in Nespor & Vogel 1986) has been widely discarded (see, for example, Anderson 2005:42–44).

3.1. ASYMMETRIES IN DISFLUENCIES. In spontaneous speech, content words are more often preceded by pauses than function words, as has repeatedly been noted in the psycholinguistics literature.⁹ What seems to have gone unnoticed, however, is the fact that this also means that pauses and other kinds of disfluencies often occur *IN BETWEEN* a preposed function word and its subsequent lexical host, as in the following example.

(5) And that's the end of **the .. story**. (Chafe 1980:308, story 9)

Importantly, the same regularity applied to the combination of a content word and a *POST*posed function word will not result in a discontinuity between the function word and its host, as seen in the following example.¹⁰

(6) Japanese (Fox et al. 1996:195; boldface added)
 ichioo zen kokumin **no (0.7)** joohoo o nigitteru, ...
 by.and.large all citizen GEN information OBJ have
 'by and large, (they) have information about all the citizens, ...'

Here the content word *joohoo* 'information' is preceded by a hesitation pause that occurs in between it and an *UNRELATED* function word (*no*) and thus does not affect the continuity between the function word and its host. What one will generally not find in Japanese discourse data is a (hesitation) pause or other kind of disfluency *BEFORE* function words such as *no* in 6.¹¹ By contrast, disfluencies in between preposed function words and their following hosts are fairly common, as the evidence gathered in this section suggests.

As this is of some importance for the current argument, a short note on the notion of 'hesitation pause' is in order. Hesitation pauses occur *WITHIN* intonation units. They are to be distinguished from boundary pauses, which belong to a bundle of features delimiting intonation units. Importantly, hesitation pauses do not involve any pitch resets in the overall intonational contour. That is, if one were to cut out a hesitation pause from a given utterance, there would be no noticeable interruption of the intonational contour when the two remaining parts are put together without the pause segment.

⁹ See, for example, Levelt (1989:302): 'Function words (articles, pronouns, and other minor category words), which are of high-frequency, were far less often preceded by pauses than content words ... Beattie and Butterworth (1979) argued that it is the *PREDICTABILITY* of a word, not its frequency, that correlates with pausing'. Note that in the current context the question of whether frequency or predictability is involved in the occurrence of disfluencies is not of import.

Statements on this issue are usually based on English data, and occasionally also on other European national languages. However, to date no one has challenged this observation, claiming that there is a language where function words are more often preceded by pauses than content words. Transcripts of spoken language from any part of the world typically confirm the observation, provided they carefully indicate disfluencies and pauses.

¹⁰ Glosses for grammatical categories are abbreviated as follows: ACC: accusative, ART: article, AV: actor voice, DAT: dative, DEM: demonstrative, DIR: direction, DIST: distal, FIN.PART: final particle, GEN: genitive, GER: gerund, HON: honorific, INJ: interjection, INSTR: instrument(al), LK: linker, LOC: locative, LV: locative voice, MED: medial (deictic), NEG: negation, NOM: nominative, OBJ: object, PL: plural, PM: predicate marker, PN: proper name, POSS: possessive, POT: potentive, PRE: preposition, PRX: proximal, PV: patient voice, RDP: reduplication, REPORT: reportative, RLS: realis, SBJ: subject, SG: singular, SPEC: specific (article), ST: stative, UG: undergoer.

¹¹ See, for example, the detailed transcriptions of spontaneous speech found in works such as Iwasaki 1993, Fox et al. 1996, Tanaka 1999, Matsumoto 2000, Hayashi 2001, 2003a,b, and Morita 2005. Note that Japanese postposed function words are not inseparable from their hosts, as shown by the phenomenon of postposition-initiated utterances in Japanese conversations discussed in Hayashi 2001. In these utterances, a second speaker interrupts and completes a construction currently under way by the current speaker, beginning with a postposition. This is a rather rare phenomenon. Hayashi found twenty-five examples in thirteen hours of conversation.

Hesitation pauses are frequently filled, either by some filler element such as *uh* or by arresting the articulators in place for the next segment.¹² Furthermore, they are typically relatively short (< 500 ms), since there is an upper limit on how long speakers are able to continue the intonational contour they had begun before the disfluency occurred. Longer hesitations may thus lead to an abandonment of the current contour and the start of a new one.

As noted above, hesitations and other kinds of disfluencies may occur after preposed function words. This involves a disruption of the prosodic word or phrase they typically form with their lexical hosts when produced fluently. What exactly is involved in such a disruption? There are two major options. The first is that the production of the prosodic word, of which the proclitic function word forms a part, is abruptly stopped. This will result in a fragment, which is an ill-formed prosodic unit. It is often represented by *th-* and *the-* in transcripts of spontaneous speech, for example, in the Pear stories and in Clark & Wasow 1998. In the second option, the preposed function word is exceptionally produced as a prosodic word of its own, which often involves lengthening and a change of vowel quality, such as when the English definite article is produced as *thee*. Clark and Wasow (1998:226) call this a PHONOLOGICAL ORPHAN. Orphans are often produced together with a hesitation word such as *uh* or *um*. While irregular in the sense that they do not conform to the regular prosodization of the function word in question, orphans are regular prosodic units in that they conform to the minimal requirements for prosodic wordhood that apply in a given language (for example, they consist of at least one foot).

As Clark and Wasow (1998:229f.; see also Levelt 1983, 1989 and Clark & Fox Tree 2002) show, fragments and orphans are not haphazard, uncontrolled productions, but reflect monitoring and planning by the speaker on various levels. Among other things, combinations of orphans and fillers may be produced as resyllabified prosodic words, as when *a uh* is pronounced as *ei.yuh* or *to uh* as *to.wuh* (Clark & Wasow 1998:229). Furthermore, the selection of *uh* vs. *um* depends on the length of the anticipated delay in further production (Clark & Fox Tree 2002). We return to this point in §4 below, where the question of why speakers actually produce preposed function words when hesitating is discussed. An alternative would be that they could suspend speech before the function word and continue only when the complete unit of preposed function word and lexical host is ready to go into production.

In the current section, our major concern pertains to the fact that hesitations involve the insertion of a prosodic boundary in between preposed function words and their hosts. In the case of fragments, the boundary is a truncation: that is, an abrupt interruption of speech, often ending in a glottal closure, followed by a pitch reset. In the case of orphans, the boundary is a prosodic word boundary. The question of whether higher-level prosodic boundaries, in particular phonological phrase and intonation unit boundaries, are possible after preposed function words is a matter that still requires further research, and that may be variable across languages. It is argued below that in Tagalog, intonation unit boundaries are possible after preposed function words. But truncations after preposed function words are probably possible and attested in many, if not all, languages that have such function words. The relevant examples evince a type of disfluency that could be called EXTREME FRAGMENTATION, that is, cases where the speaker produces only a word

¹² The phonetic description of hesitations here follows Simpson 2006 and the work referred to there, in particular Local & Kelly 1986.

or short phrase at a time and then pauses, before continuing the clause or sentence under construction. The following is a typical example from German.

- (7) German (SELTSAM 18ff.; author's corpus)
 un: dann hab' ich plötzlich von weitem,
 and then have I suddenly from afar
 (0.5) gesehen **dass**,
 seen that
 (0.8) en Teil **von der** –
 a part of the
 (0.8) öh öh (0.4) Strecke,
 uh uh road
 öh mit Schnee,
 uh with snow
 öh (0.8) ähm (1.1)
 uh um
 also (0.4) mit Schnee bedeckt war.
 well with snow covered was

There are two points of interest in this type of example. First, preposed function words are often chunked with the preceding word rather than with their morphosyntactic host, which follows the discontinuity. Thus, in 7 the complementizer *dass* is produced right after the matrix verb and thus is separated from the subordinate clause it introduces. And the combination of preposition plus article *von der* is separated from the noun *Strecke* with which they form a PP (and a DP). Second, it is arguably the case that examples such as these involve a rapid succession of intonation unit boundaries, both regular and truncated, as indicated in the transcript (each line equals one intonation unit).

The separation of preposed function words from their following hosts in disfluencies can be of relevance for the suffixing preference only if it is not a rare, sporadic phenomenon but occurs with sufficient frequency in spontaneous discourse. It is admittedly not a straightforward task to define what 'sufficient frequency' means in this context. The following data, however, should suffice to give some plausibility to the hypothesis that this phenomenon is indeed of relevance for the suffixing preference. This is especially the case if one keeps in mind that no such separations occur with postposed function words and their preceding hosts, a topic to which we attend further below.

Table 4 provides data on the disfluencies occurring for five frequent preposed function words in the Pear story corpus (Chafe 1980). This corpus consists of retellings of a short film and has a fairly high number of disfluencies, as the speakers try to remember what they saw and find the appropriate words to express it. The narratives are segmented into intonation units (1,989 units altogether; see Croft 1995 for further details on prosodic and grammatical units in this corpus). The transcription systematically notes all disfluencies and pauses, including in particular very short breaks in continuity (represented by ..).

The first row in Table 4 gives the total number of function words in the whole corpus, with the definite article *the* being by far the most frequent word, unsurprisingly.¹³ The

¹³ The numbers for *of* in this and the following tables exclude occurrences in the hedging expressions *sort of* and *kind of* (e.g. *he sort of turns*, *sort of like southern California*, *spoken language kind of communication*, etc.). These function as units and, of course, typically also involve disfluencies. The numbers for disfluent *of* would be distinctly higher if these cases were included. The numbers for *have* are for uses as a temporal auxiliary (forms: *have*, *has*, *had*) with a following participle (*has clambered down*), not to uses as a possessive

	<i>the/thee</i>	<i>of</i>	<i>in</i>	<i>have</i>	<i>be</i>
TOTAL	984	237	165	29	85
disfluency IN BETWEEN FW and host	82 (8.3%)	22 ^a (9.3%)	25 ^b (15.1%)	6 (20.7%)	17 (20.0%)
disfluency only BEFORE FW and host	36 (3.6%)	4 (1.9%)	5 (3.0%)	1 (3.4%)	5 (5.9%)
FW repeated after disfluency	35 (3.6%)	5 (2.1%)	6 (3.6%)	—	1 (1.2%)

TABLE 4. Frequency of disfluencies involving three preposed function words (FWs) in the English Pear stories (Chafe 1980, appendix).

^a Includes twelve instances where *of* is followed by an article before the hesitation as in [3.55[.6] *And then u—m [2.1]] you hear the sound of a .. sheep* (Chafe 1980:310, story 11).

^b Includes eleven instances where *in* is followed by an article before the hesitation as in example 11.

second row gives the number of instances where these function words have been separated from their syntactic hosts by a disfluency, as in the following example.

(8) .. and he [1.25] he wipes **the** .. dirt off of him, (Chafe 1980:311, story 11)

In rare instances, the function word is both preceded and followed by a hesitation pause.

(9) like the movie [.2] **the** .. sound track had been slowed down,
(Chafe 1980:314, story 14)

The hesitation sometimes consists of a false start, as in 10.

(10) [.35? and [.2]] there's a rock in **the r** road, (Chafe 1980:305, story 5)

There are also instances where the disfluency involves two preposed function words, as in 11.

(11) [.75] And he kneels down to put the pears **in the** .. third empty basket,
(Chafe 1980:314, story 14)

It should be noted that here only hesitation pauses within intonation units are of interest, not the pauses occurring in between two units, which in these transcripts are put at the beginning of a line (e.g. the pause of .75 seconds at the beginning of example 11 is a boundary or planning pause, not a hesitation pause). Boundary pauses are not included in any of the numbers presented here.

The third row pertains to instances where the disfluency occurs before the function words only, that is, where the function word and its host are not separated from each other by the disfluency, as in 12.

(12) And he sees .. **the** pears, (Chafe 1980:305, story 5)

As the numbers in Table 4 and the following tables make clear, these instances are significantly less common than those where function word and host are separated by the disfluency.

The last row represents instances where the function word is repeated after the hesitation pause or false start as in 13 and 14, which contains a sequence of function words repeated after the disfluency. The totals given in this row pertain to the number of examples where such a 'doubling' of function words occurs (i.e. the function words are not counted twice).

predicate (*had three baskets*). The numbers for *be* are for uses as a temporal auxiliary (forms: *is, are, was, were*) with following participle (*was picking pears*), not to uses as a copula (*is important*) or existential predicate (*there were two baskets*). The frequent clitic form *'s* (as in *he's going down the road*) has not been included in this count since it is unclear how disfluencies before the phonological host (i.e. *he* in the example just given) would have to be classified.

- (13) but he puts **the .. the** big basket of pears on his bicycle, (Chafe 1980:308, story 9)
 (14) [35] And it shows him sort of tumbling them into a .. **in a** [.4] **in a** basket,
 (Chafe 1980:313, story 13)

As pointed out by Kie Zuraw (p.c.), these instances are ambivalent with regard to their relevance for the transition from clitic to affixal status. On the one hand, they increase the number of instances where a function word is involved in a disfluency and is separated from its host by it. On the other hand, the fact that the function word is repeated after the disfluency could be interpreted as reflecting a particularly close bond between function word and lexical host, and thus promoting the dependence between them.

Other corpora show very similar figures. Table 5 presents the numbers for twenty-four German Pear stories, also recorded and transcribed by Wallace Chafe and colleagues.¹⁴ They consist of 1,532 intonation units and include a few very short retellings. The table provides data for two case forms of the German definite article, and the third singular and plural forms of three auxiliaries: *ist/sind* 'is/are', *hat/haben* 'has/have', and *wird/werden* 'will(SG/PL)'. The auxiliaries occur in periphrastic tense forms (e.g. *ist gegangen* 'has gone', *haben gezeigt* 'have shown') and passive formations (*wird gezeigt* 'is shown'). They have been combined into a single column AUX because the frequencies for each form are quite low. Because of extensive syncretism in the article paradigm, the article forms are multifunctional: *der* is not only masculine nominative singular (*der Mann* 'the man'), but also feminine genitive singular (*der Frau* 'the woman's') and genitive plural in all genders (*der Männer* 'the men's', *der Frauen* 'the women's', *der Kinder* 'the children's'). These forms also function as personal and relative pronouns, but, of course, have not been counted when occurring in these functions.

	<i>der</i>	<i>die</i>	AUX
TOTAL	245	249	135
disfluency IN BETWEEN FW and host	23 (9.4%)	23 (9.2%)	11 (8.1%)
disfluency only BEFORE FW and host	3 (1.2%)	12 (4.8%)	1 (0.7%)
FW repeated after disfluency	7 (2.9%)	6 (2.4%)	1 (0.7%)

TABLE 5. Frequency of disfluencies involving two forms of the definite article and three auxiliaries in the German Pear stories collected by Wallace Chafe and colleagues (unpublished ms.).

Note that auxiliaries in German may occur both before and after the main verb, as in *wird gezeigt* and *gezeigt wird* 'is shown', respectively. Both orders are well attested in the corpus (seventy-five AUX V, sixty V AUX). Disfluencies between auxiliary and verb occur only when the order is AUX V.

That the preceding data are not totally haphazard and just a characteristic of retellings of the Pear story can be clearly seen in Table 6, which presents data from the Santa Barbara Corpus of Spoken American English (SBC). This corpus consists of approximately 249,000 words and 70,000 intonation units, and '[t]he predominant form of language use represented is face-to-face conversation, but the corpus also documents many other ways that that people use language in their everyday lives: telephone conversations, card

¹⁴ This set of Pear stories has not been published to date. I am very grateful indeed to Wallace Chafe for making the transcript available to me. My own set of Pear stories, which have been transcribed according to the same guidelines as the ones used by Chafe and colleagues (i.e. intonation units as the primary segments, inclusion of all hesitations, even very short ones), shows similar numbers. The corpus is much smaller (only 594 intonation units), though. The form *der* is attested seventy-nine times altogether, with seven tokens (8.9%) followed by a disfluency, and one (1.2%) where the disfluency occurs only before the article. The form *die* is found seventy-four times altogether, with six tokens (8.1%) followed by a disfluency, and one (1.3%) preceded by one.

games, food preparation, on-the-job talk, classroom lectures, sermons, story-telling, town hall meetings, tour-guide spiels, and more'.¹⁵ Importantly for our purposes, the transcription details are basically the same as the ones used in the earlier work by Chafe and colleagues.¹⁶

	<i>the</i>	<i>of</i>	<i>in</i>
TOTAL	8,320	3,235	2,824
disfluency IN BETWEEN FW and host	624 (7.5%)	196 (6.1%)	188 (6.6%)
disfluency only BEFORE FW and host	110 (1.3%)	44 (1.4%)	47 (1.7%)
FW repeated after disfluency	53 (0.6%)	4 (0.1%)	12 (0.4%)

TABLE 6. Frequency of disfluencies involving three preposed function words in the Santa Barbara Corpus of Spoken American English (Du Bois et al. 2000, 2003, Du Bois & Englebreton 2004, 2005).

The preceding observations and data are of potential relevance to the suffixing preference, because similar phenomena do not occur with postposed function words. In languages like Japanese that make heavy use of postposed function words, disfluencies also regularly occur AFTER the postposed function word, and hence typically after a phrasal constituent and not within it, as already illustrated in example 6 at the beginning of this section. The following is an example of extreme fragmentation in Japanese, in parallel with the German example in 7 above.

(15) Japanese

(Matsumoto 2000:520; boldface added)

Yamato san **ga ne:**,
 Yamato HON NOM FIN.PART
 'Mr. Yamato'
 kekkyoku **ne:**,
 in short FIN.PART
 'in short'
 wareware **o ne:**,
 us ACC FIN.PART
 'us'
 sofuto **no ne:**,
 software GEN FIN.PART
 'software'
 shigoto ni taisuru.
 job to toward
 'to the job'
 hyooka ga ano hito **ne:**,
 evaluation ga that person FIN.PART
 'evaluation, he'

¹⁵ Quote from <http://www.linguistics.ucsb.edu/research/santa-barbara-corpus>, the first page of the corpus.

¹⁶ The British National Corpus (BNC) provides less detail with regard to hesitation phenomena. Consequently, the numbers there are considerably lower. Thus, for example, in the spoken conversation part of the BNC (4,233,962 words) we find 115,141 tokens of *the* with the part-of-speech label ART. Of these, 3,595 (3.1%) are found before a pause or the hesitation particles *erm*, *er*, or *eh*. Thus, the overall frequency of definite articles separated from their host by a disfluency appears to be clearly lower than in the data looked at so far. However, the difference is in all likelihood primarily due to the fact that the definition of a 'pause' in the BNC differs significantly from that used in the Pear story and SBC data. While the transcriptions of the Pear stories and the SBC include very short hesitations ('a break in timing too short to be measured as pause'; Chafe 1980:301), the BNC indicates only 'significant pauses', which are defined as 'silence, within or between utterances, longer than was judged normal for the speaker or speakers' (<http://www.natcorp.ox.ac.uk/docs/URG/cdifsp.html#cdifsp>). And while the Pear stories include measured pauses from 100 milliseconds and up, in the BNC pause duration 'is normally specified only if it is greater than 5 seconds' (ibid.).

shitenai.
do.NEG
‘doesn’t do’
.. hyooka **o**.
evaluation ACC
‘(high) evaluation’
.. zettai.
never
‘never’

As in the German example above, the prosodic discontinuities here typically occur after function words (the final particle *ne*: should be considered part of the prosodic packaging, being a regular part of the boundary marking of intonation units). But unlike in the German example, this does not lead to discontinuity in between function words and their lexical hosts. In fact, the hesitations occur at regular prosodic boundaries (word or phrase). What is disrupted here is the continuity between the constituents of a higher-level unit, that is, a clause.¹⁷

The different prosodic treatment of function words is found not only across languages—where there may always be some doubt as to the comparability of the data—but also within a language (and within the same text produced by the same speaker, for that matter). A language that makes use of both pre- and postposed function words with sufficient frequency is Tagalog. High-frequency preposed function words are the phrase markers shown in Table 7.

COMMON NOUNS		PERSONAL NAMES	
ang	‘specific article’	si	‘personal article’
ng [naŋ]	‘genitive’	ni	‘possessive’
sa	‘locative’	kay	‘dative’

TABLE 7. Tagalog phrase markers.

These come in two paradigms, one for personal names and one for all other object expressions. The distribution of these two sets of phrase markers differs slightly, as indicated by the different glosses, but this is of no further relevance in the current context. What is of relevance is the fact that these elements clearly are proclitics, occurring at the left edge of the phrase they mark. Simple examples are *ang=bahay* ‘the/a house’ and *si=Pedro* ‘Peter’. They cannot be stressed and cannot host other clitics, as illustrated in 16.

- (16) sa Manila **pa** ‘still in Manila’
*sa **pa** Manila

The negative marker *hindi*’, by contrast, is clearly not a clitic element, since it allows stress and can host second-position clitics.

- (17) Tagalog (mayon 101)¹⁸
hindi **pa** silá nakàkalayó’,
hindi’ pa silá naka.RDP-layo’
NEG still 3PL POT.AV.RLS.RDP-distance
‘they hadn’t gotten far yet,’

¹⁷ Matsumoto observes with regard to example 15: ‘It should be noted that speaker A could have conveyed the same ideational content as the 9 short IUs collectively do by means of one full-clausal IU, which could roughly correspond in English to “Mr. Yamato never values our software job”’ (2000:521).

¹⁸ Tagalog examples are taken from my corpus of spoken Tagalog narratives and quoted with reference to the text identifier and line number.

The enclitic *pa* belongs to a larger set of second-position clitics that are phonologically enclitics. Example 18 lists the high(er)-frequency enclitics, which include personal pronouns such as *siyá* (the third-person singular *ang*-form, usually pronounced [ʃja] or [ʃa]) and *niyá* (the corresponding *ng*-form, usually pronounced [nja]).

(18) Tagalog enclitics (second-position clitics)

<i>na</i>	completive ('already'); linking element in between constituents of a phrase
<i>pa</i>	incompletive ('still')
<i>lang/lamang</i>	'only'
<i>ba</i>	question
<i>siyá, niyá, silá, nilá</i> , etc.	clitic personal pronouns

The preposed phrase markers are frequently separated from their lexical hosts by disfluencies, just like articles and prepositions in English or German. The most common type of prosodic discontinuity, resulting in phonological orphans, consists of a long and level lengthening of the vowel, followed by a pause, as in 19.

(19) Tagalog

(mayon 010)

at itú-ng **si::** (0.3) Magayón na itó,
and PRX-LK PN Magayon LK PRX
'and this Magayon,'

As seen in the waveform and F0-extraction for this example given in Figure 1, the lengthening is very pronounced, longer than half a second and thus longer than the following pause. Intonation on the lengthened segment is basically level, the overall intonational contour being continued after the pause. Also note the fairly strong fall on the final syllables, which is typical for one very common type of continuing intonation unit boundary. This will be of relevance shortly.

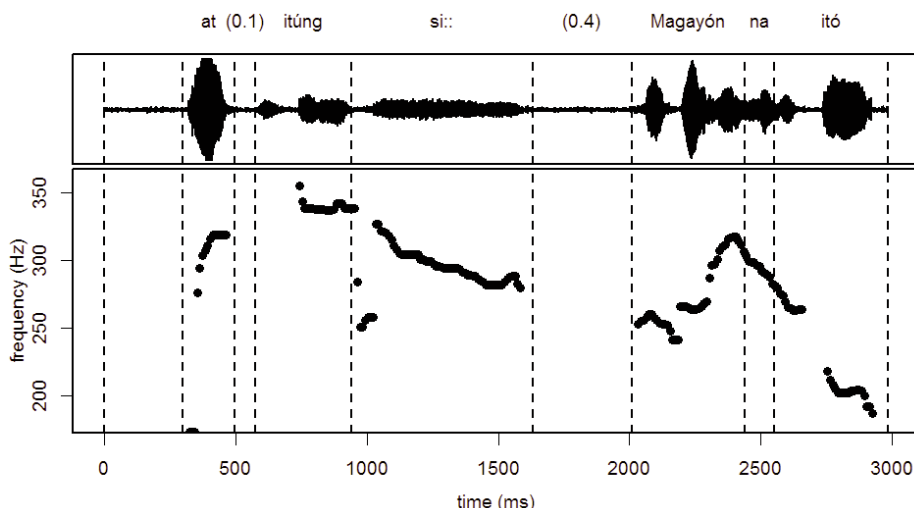


FIGURE 1. Waveform and F0-extraction for example 19.

As Streeck (1996) has noted, lengthening is the typical form of hesitations in many insular Southeast Asian languages that generally lack hesitation particles such as *uh* and its cognates in other languages. There are, in fact, no hesitation particles in the author's spoken Tagalog corpus.

Unlike in the case of English and German, where discontinuities between function words and their hosts generally do not involve higher-level prosodic boundaries, Tagalog phrase markers may be separated from their hosts by an intonation unit boundary. Both *sa* and *na* in example 20 are lengthened, but the lengthening here is much shorter than in example 19, and thus typical for unit-final lengthening. Furthermore, the fall in pitch on *sa* and *na* is abrupt and steep, as it is on the final syllables of example 19. Most importantly, there is a pitch reset in the subsequent unit that provides clear evidence that we are dealing with an intonation unit boundary here; see Figure 2.

(20) Tagalog (pep 001-3)

ang akíng kwento ay tungkól **sa::**,

SPEC 1SG.DAT:LK story PM about LOC

(0.2) isá-ng pangyayari **na::**,

one-LK event LK

(0.8) <kə> may kasaysayan sa isá-ng manlalakbáy.

EXIST statement LOC one-LK traveler

‘My story is about, one incident that, is told about a traveler.’

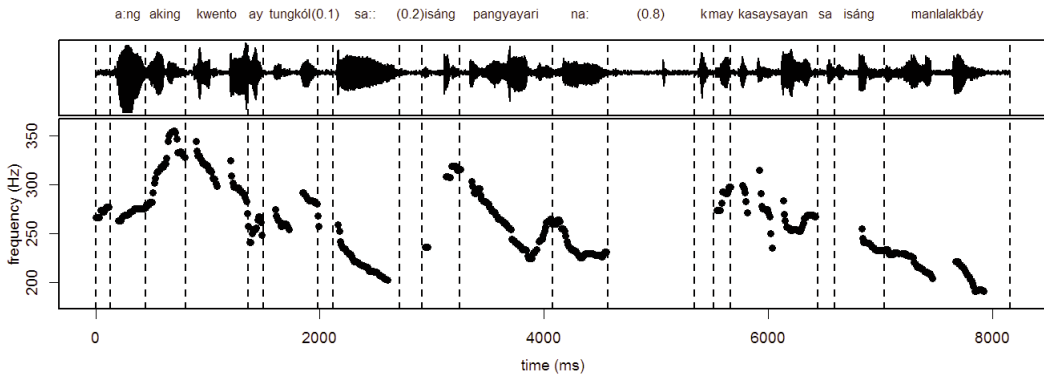


FIGURE 2. Waveform and F0-extraction for example 20.

In example 21, the evidence for the preboundary placement of *sa* is perhaps even stronger. As seen in Figure 3, the phrase marker *sa* at the end of the first line is minimally lengthened and there is a clearly audible breath intake in the following pause. Intonation falls sharply on *sa* and there is a clear pitch reset on *lugár*. Hence, this is not a unit-internal disfluency, but a regular chunking of information across two intonation units. The discontinuity after genitive *ng* in the second line, by contrast, shows all of the tell-tale signs of a hesitation disfluency.

(21) Tagalog (pep 007FF)

o: isá-ng manlalakbáy **sa**,

or one-LK traveler LOC

(0.3) lugár **ng::** (1.4) Laguna.

place GEN

‘(Pepito was a stranger) or a traveler in, the area of, u::m, Laguna.’

The boundary at the end of the first line in example 21 is the same type of intonation unit boundary that commonly occurs on postposed clitics, as is seen in example 22 on the enclitic completive marker *na*. See Figure 4 for details.

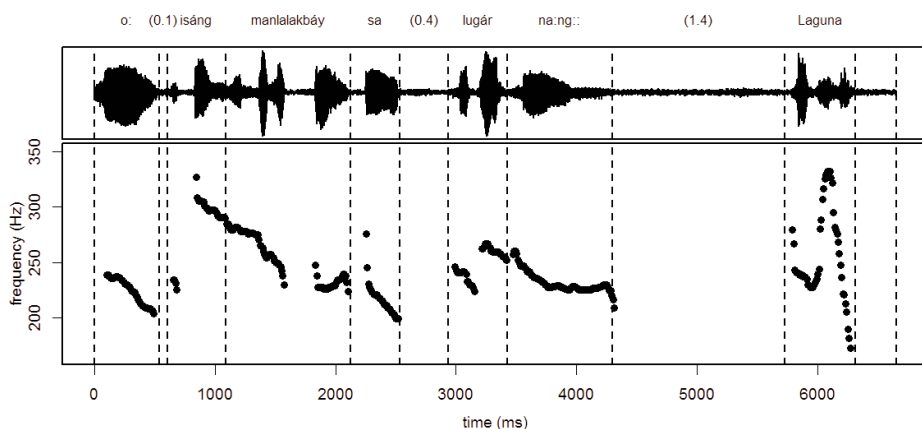


FIGURE 3. Waveform and F0-extraction for example 21.

(22) Tagalog

(lamok 040)

nung maka-tulog **na**,
 DIST.GEN:LK POT.AV-sleep already
 (0.7) itú-ng kanyá-ng hinihingutuhan.
 PRX-LK 3SG.DAT-LK RLS.RDP:delouse:LV
 ‘Then (she) fell asleep, the one she was delousing.’

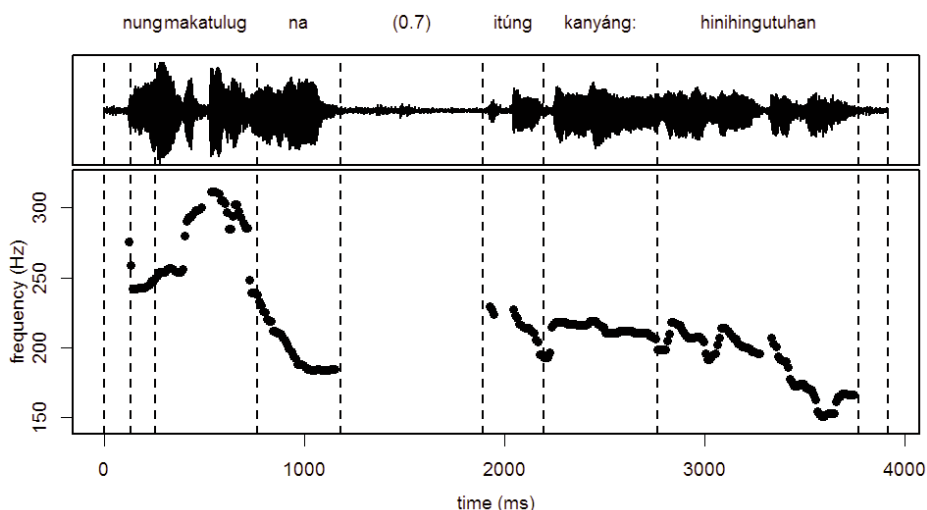


FIGURE 4. Waveform and F0-extraction for example 22.

In the case of enclitic function words, it is to be expected that they may occur at intonation unit boundaries and carry a boundary tone, because they typically occur at the end of prosodic words and phrases. For proclitic function words, this is the exception rather than the rule. Still, it appears to be possible to separate preposed function words from their hosts by an intonation unit boundary in Tagalog, as the preceding examples have shown. Importantly, the converse is not attested; there are NO enclitic function words that are separated from their hosts by a disfluency or by regular prosodic phrasing. That is, neither of the following two hypothetical options is an attested possibility for rendering example 22.

- (23) *nung makatulo:g (0.7) na itung kanyang hinihingutúhan.
 *nung makatulog, (0.2) na itung kanyang hinihingutúhan.

In the first option, a hesitation occurs between enclitic *na* and its host *makatulog* (with lengthening of the final syllable of *makatulog*, as typical for hesitations). In the second option, a regular intonation unit boundary is inserted between the two.

Table 8 provides data on the frequency of prosodic discontinuities after preposed function words in the Tagalog corpus of spontaneous narratives available to the author. This corpus consists of 1,297 intonation units.¹⁹

MARKER	TOTAL	before disfluency or IU boundary (% of total)	repeat after discontinuity
ang	198	12 (6.1%)	1
ng	127	17 (13.4%)	0
sa	168	31 (18.4%)	7
si	22	6 (27.3%)	0
mangá	75	15 (20%)	0

TABLE 8. Frequency of prosodic discontinuities after preposed function words in the author's Tagalog narrative corpus.

There are no instances of preposed function words AFTER a prosodic discontinuity, with the exception of the ones repeated after such a discontinuity (listed in the fourth column of Table 8). There is not a single instance of an enclitic function word, including the enclitic forms of the pronouns, that is separated from its host by a prosodic discontinuity.²⁰ That is, there is a very clear asymmetry between pre- and postposed function words in this regard, as predicted by the hypothesis in 4.

The data reviewed in this section, while certainly in need of further expansion and corroboration, make it clear that prosodic discontinuities after preposed function words are not sporadic phenomena, but evince systematic features and occur with sufficient frequency to be of potential relevance in explaining the suffixing preference. Before this issue is explored further in §4 below, we first briefly take a closer look at the related asymmetry in phonological attachment options for proclitics and enclitics.

3.2. THE CLITIC ASYMMETRY. The asymmetry between preposed and postposed function words with regard to prosodic boundaries is reflected on a more grammaticized level in an asymmetry between proclitics and enclitics. Elements that morphosyntactically are clearly proclitics may regularly be enclitics on the phonological level: that is, they are phonologically attached to a structurally unrelated host. Perhaps the most well-known example of this kind are determiners and prepositions in K^wak^wwala (also known as Kwakiutl).²¹

¹⁹ *mangá* (in standard orthography rendered as *mga*) in the last row of this table is a proclitic nominal plural marker that typically occurs immediately after the phrase marker, as in *marami sa mangá alagád ko* (many LOC.PL follower 1SG.POSS) 'many among my followers'. It has been added here to show that the phenomenon under discussion is not limited to the phrase-marking proclitics listed in Table 7.

²⁰ There are 155 tokens of enclitic *na* and thirty-one of enclitic *pa* in the corpus, hence enough instances for disfluencies or IU boundaries to occur. In fact, forty-seven out of the 155 tokens of *na* and three out of the thirty-one tokens of *pa* precede a disfluency or IU boundary. But, as just noted, there is never a disfluency preceding them.

²¹ Morpheme segmentation and glosses follow the quoted source. However, hyphens have been replaced by equal signs in line with the widely used convention for representing clitic boundaries. See Anderson 2005: 14–22, 101–6, *passim*, for a fuller discussion of K^wak^wwala clitics. Fortescue (2006) provides a historical scenario of how the Northern Wakashan system came about, which is absent from Southern Wakashan.

- (24) K^wak^wwala (Anderson 1992:18)
 nanaqəsil=**ida** iʔgəl'wat=i əliwinux^wa=**s=****is** mestuwi la=**xa**
 guides=**SBJ/ART** expert=**DEM** hunter=**INSTR**=**his** harpoon **PRE=OBJ/ART**
 mig^wat=i
 seal=**DIST**
 'An expert hunter guides the seal with his harpoon.'

Basically, every prosodic word in this example includes one or two enclitics that morphosyntactically belong to the following lexical item, the exception being the final word where the phonological enclitic is also a morphosyntactic one.²² Thus, for example, the instrumental 'preposition' =s and the possessive pronoun =is in 24 form a phonological unit with the preceding lexical base əliwinux^wa 'hunter' despite the fact that morphosyntactically they form a PP with mestuwi 'harpoon'. Figure 5 from Anderson 1992 nicely illustrates the mismatches between the two levels.

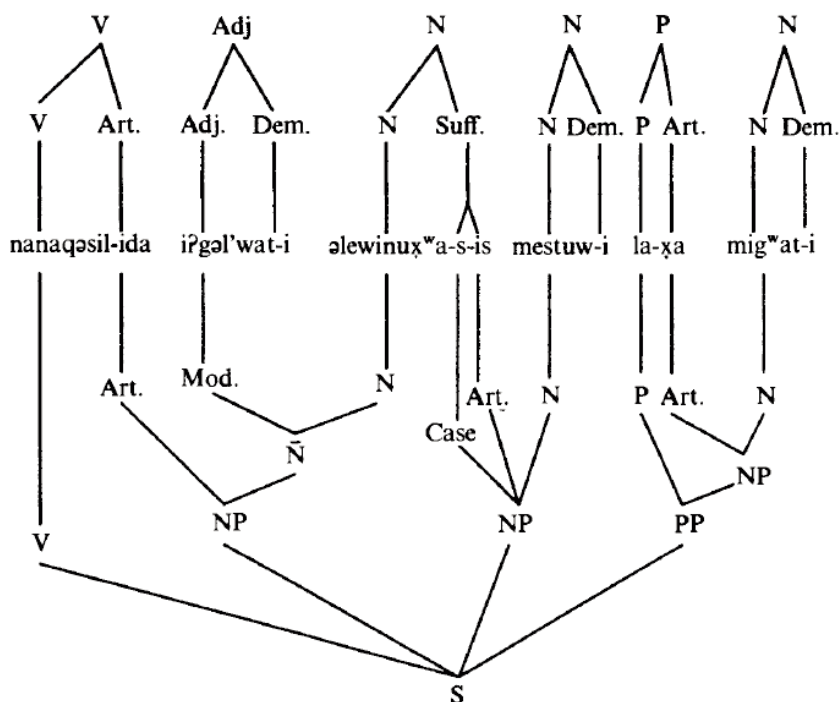


FIGURE 5. Phonological and syntactic organization of a K^wak^wwala sentence
 (= figure 1.5 from Anderson 1992:20; reproduced with permission).

Similar phenomena are found elsewhere in the world's languages, including some Philippine languages. The following examples from Central Bontok contain an element =s that in all likelihood is a cognate of the Tagalog proclitic locative phrase marker *sa* discussed in the preceding section. Unlike Tagalog *sa*, the Bontok locative marker REGULARLY phrases with the preceding prosodic word. As in the K^wak^wwala examples, the

²² A further exception is the demonstrative enclitic to iʔgəl'wat 'expert'. As Anderson (2005:104f.) illustrates, the clitic =i here and in final position belongs to a set of 'postnominal' second-position clitics that differ distinctly from the preposed clitics of major concern here.

clitic attaches phonologically to whatever happens to precede it, regardless of the category and the constituency of the preceding item. In 25 = *s* attaches to the main predicate (after the agent pronoun, which is also an enclitic); in 26 it is found at the end of a noun phrase, attaching to the preceding noun.²³

- (25) Central Bontok (Reid 1970:38)

balód-en=da nan toló=way p-in-ásek ay p-in-ásek=na=s ábong=da
bind-PV=3PL SPEC three=LK RLS-split LK RLS-split=3SG=LOC house=3PL
ya eney=na=s an da asáwa=na
and carry=3SG=LOC place PL wife=3SG

‘(the man and his helper) bind three split logs that he had split at their house and then he takes it to his wife’s place’

- (26) Central Bontok (Reid 1970:23)

in-manok nan babái=s nan masdem
AV-chicken SPEC woman=LOC SPEC night

‘the woman performs a chicken sacrifice in the evening’

Further examples of a mismatch between phonological and morphosyntactic hosts in cliticization are reviewed by Cysouw (2005), who presents a total of ten examples from Cape York languages (Pama-Nyungan), Djinang/Djinba (Pama-Nyungan, Arnhem Land), Kherwarian languages (North Munda), Udi (Lezgian), Ingush (Nakh), Northern Talysh (Iranian), Greek (Ancient and Modern), Northern Mansi (Uralic), and Yagua (Peba-Yaguan), in addition to K^wak^wwala. His survey does not include Bontok and Yucatec Maya, for which Lehmann (1995:133) notes that ‘subject pronouns precede the verb, and possessive pronouns precede the possessed noun; but both are enclitic to whatever happens to precede them’.

Clitics showing these placement characteristics have been called DITROPIC CLITICS (Cysouw 2005:18, Spencer & Luís 2012:66). They include the well-known examples of fusions of prepositions with articles (e.g. French *du* (< *de le*) or German *vom* (< *von dem*)) and reduced auxiliaries as in English *They’ve seen it* or *You’ll see*. These examples also show the characteristic mismatch between morphosyntactic [*von* [*dem* N]_{DP}]_{PP} and phonological [*vom*]_{p-word} [N]_{p-word} constituency. It is less widely recognized, however, that, even in European national languages, ditropic clitic placement is found to be very similar to that seen in the K^wak^wwala and Bontok examples above. Gussenhoven and Jacobs (1998:250f.), for example, claim that function words in Dutch are generally included in the prosodic word to their left, if there is one (exceptions being function words such as *of* ‘or’ and *en* ‘and’, which are given prosodic word status in the lexicon). They illustrate this with examples in which a clitic complex consisting of a preposition and an article cliticizes to the preceding verb.²⁴

- (27) Dutch (Gussenhoven & Jacobs 1998:251)

het _ω(χa:t.ti.nət)_ω putje
it goes.into.the drain
‘it goes into the drain’ (cf. [χa:t]_V [ɪn]_{Prep} [ət]_{Art})

For current purposes, of interest are all instances where a preposed function word (which of course excludes second-position clitics) cliticizes to a structurally unrelated preceding host. This includes in particular the case of reduced auxiliaries and other left-

²³ The following Bontok examples differ in minor detail from the published versions. Corrections were provided by the author, Lawrence Reid, via an email exchange.

²⁴ Spencer (1991:376) provides related examples from English, such as *Put [tʃə] hands up* ‘put your hands up’ and *Giv [ntə] me* ‘give them to me’.

attaching function words, as in the Dutch and English examples just given. Clearly, then, there is ample evidence that preposed function words may form part of a phonological unit that precedes the unit they belong to morphosyntactically.

What is of major import here is the fact that the converse phenomenon for postposed function words—that is, postposed proclitics—does not appear to be attested (see also Cysouw 2005:18). This statement is in contradiction to the clitic-positioning claims proposed in Klavans's (1985) classic paper on syntax/phonology mismatches in cliticization and thus requires further substantiation. Klavans claims that the positioning of clitics is governed by three binary parameters, given in 28.

- (28) Parameters for clitic placement (Klavans 1985:97f.)
- a. Dominance: INITIAL/FINAL in its domain (S, VP, NP, etc.)
 - b. Precedence: BEFORE/AFTER its (morphosyntactic) host
 - c. Phonological liaison: PROCLITIC/ENCLITIC

These parameters provide for the eight possibilities of clitic placement shown in Table 9.

TYPE	PARAMETER VALUES	TYPE	PARAMETER VALUES
1	initial/before/enclitic	8	final/after/proclitic
2	initial/before/proclitic	7	final/after/enclitic
3	initial/after/enclitic	6	final/before/proclitic
4	initial/after/proclitic	5	final/before/enclitic

TABLE 9. Types of clitic placement according to Klavans 1985:103.

The four types of major interest here are type 1 and its mirror-image, type 8, and type 4 and its mirror-image, type 5 (all given in boldface in Table 9). Types 1 and 5 are well attested in K^wak^wwala and other languages, as just illustrated. For type 8, however, Klavans only provides examples from Classical Greek, which is quite problematic. In Classical Greek, proclitics can occasionally be positioned AFTER their host, as shown by the examples in 29.

- (29) Classical Greek (Klavans 1985:110)
- ek Spártēs 'out of Sparta'
- kakōn éks 'out of (the) evils'

There are two problems with these examples. First, the fact that the presumed clitic receives an accent of its own when it is not in prehost position raises the question of whether it is still a clitic when it occurs phrase-finally. Second, there is no evidence for the claim that it is PROclitic in final position, one problem being to determine what it should be proclitic to, as it occurs in phrase-final, typically prepausal position.

With regard to the other potential counterexample to the claim that postposed function words cannot be realized as proclitics, that is, type 4, Klavans also provides just one, highly speculative example from Tepecano where this type only occurs in a hypothesized intermediary stage in a historical development from type 2 to type 3 (Klavans 1985:113f.). Cysouw (2005:20), in summarizing a brief review of Klavans's critics, concludes: 'Klavans gives examples of all four ditropic clitic types, though most of her examples are not very convincing or [are] factually doubtful, as has been repeatedly noted by commentators'. Only her examples for types 1 and 5 'turned out to be factually irrefutable cases of ditropic clitics'.²⁵

To conclude this section, there is ample evidence for the claim that in some languages preposed function words regularly form a prosodic unit with a structurally unre-

²⁵ For a similar assessment, see Spencer & Luís 2012:45–47.

lated host, and thus are regularly separated from their morphosyntactic host by a prosodic boundary. This prosodic boundary presumably plays a significant, if not the major, role in preventing these preposed function words from becoming prefixes.²⁶ Since the converse case for postposed function words is not attested, this asymmetry between preposed and postposed function words may be seen as a contributing factor to the suffixing preference. It is difficult to gauge how significant a factor it is, however, because the relevant information is lacking for many languages. I would not want to exclude the possibility that at least a subset of preposed function words in the languages that have them are regularly phrased with a structurally unrelated preceding host, as we have seen here in examples from K^{wak}w^wala, Bontok, English, and Dutch. But this requires further empirical research.

Regardless of whether 'misaligned' function words turn out to be a much more widespread phenomenon than acknowledged to date, it would of course not provide a full explanation of the suffixing preference, because it would simply shift the explanandum to another level. The question raised by the data reviewed in this section is this: Why are preposed function words separated from their lexical hosts more often by a prosodic boundary than postposed function words? What we are looking for here are factors that underlie both of the 'optional' discontinuities caused by disfluencies discussed in §3.1 and the more grammaticized prosodic phrasing seen in §3.2. The following section proposes two such factors, both pertaining to the on-line production of spontaneous speech.

4. FUNCTION WORDS, PROCESSING, AND TURN-TAKING. This section reviews two mutually reinforcing explanations for the asymmetries discussed in the preceding section, one from language processing and one from turn-taking, and tries to show how they may also be relevant for explaining the suffixing preference.

There is substantial evidence in the speech production literature that function words (closed-class items) are processed somewhat differently from content words. Within the basic architecture of speech production developed in Levelt 1989 and Levelt et al. 1999, one difference pertains to the level of lexical selection. The selection of function words is part of grammatical encoding and thus at least partially driven by the syntax rather than by the semantics (the propositional content to be conveyed), although semantics often also plays a role. More importantly in the present context, many function words are high-frequency items and—in context—usually also highly predictable. High-frequency items are supposed to have a lower activation threshold (Levelt 1989: 203 and further references there) and are thus 'ready to go', as it were, once selected in the grammatical encoding process. Additionally, function words, like all lexical items, are subject to the local repetition effect (or, in Levelt's (1992) terms, the IMITATION PROBLEM), which in turn involves a recency effect. That is, there is a general tendency for interlocutors to use recently used lexical items and syntactic constructions again and again (Levelt 1992:8). This may be caused 'by a temporary extra activation of the relevant lemma, due to the speaker's hearing or using the word' (Levelt 1992:8). The ready availability of function words is relevant for the asymmetry in prosodic phrasing between preposed and postposed function words in the following way.

The Levelt model assumes that phonological encoding involves two independent steps, namely, the generation of a metrical frame (the metrical skeleton of a prosodic

²⁶ The question of whether it is at all possible that such 'wrongly' attached function words could ever become true suffixes to their phonological hosts and what factors prevent this from happening cannot be further explored here. Most probably, the factor of (lacking) semantic relevance invoked in the work of Joan Bybee (including Bybee et al. 1990) is of major import in this regard.

word) and the filling of the slots of this frame with the phonological segments associated with the lemmata that were previously selected at the level of grammatical encoding. The separation of these processes is necessary in order to achieve the correct syllabification of spoken words in context. That is, while *escort* as a prosodic word on its own is syllabified as *e.scort*, it is syllabified as *e.scor.tus* when forming a prosodic word with clitic *us*.²⁷ As this example shows, clitic function words are of major import for this feature of the overall model (the other major phenomenon of relevance in this regard is derivational morphology).

The level of phonological encoding is followed by phonetic encoding, where the gestural score for the articulator is being computed. This is done on a syllable-by-syllable basis and, importantly, in an incremental fashion:

syllabification can start on the initial segments of a word without having all of its segments. Only initial segments and, for some words, the metrical structure are needed to make a successful start. When given partial information, computations are completed as far as possible, after which they are put on hold. (Levelt et al. 1999:23)

Assuming this model of speech production, hesitations after preposed function words arise when the spelling out of the syllables of a given prosodic word comes to a hold, for whatever reason (e.g. problems in segment retrieval or in implementing the articulation of the next syllable, replanning of the overall message or the grammatical structure, on-the-fly changes in lemma selection, etc.). Given the high activation state of function words, it is very rare that problems arise in spelling them out. Hence they always can go into production as soon as they are called upon. For preposed function words, this means that their articulation often follows immediately upon conclusion of the preceding prosodic word. Problems with regard to the following content word may then lead to a prosodic discontinuity between function and content word, the function word being produced either as a fragment or as an orphan (see §3.1 above).

Hesitations before postposed function words are very rare or nonexistent for the same reason that they are common after preposed ones: since postposed function words are also high-frequency items showing a high level of activation, they are readily produced as soon as the last syllable of the preceding content word is finished. In the case of enclitics, they are actually part of the metrical skeleton of their lexical host, and there is no reason to expect any disfluencies when the execution of the articulatory score for the prosodic word has already progressed to its final constituent(s). The asymmetries between the two positional types of function words with regard to disfluencies thus basically follow from the incremental left-to-right nature of speech production.

However, this cannot be the whole story, for the following reasons. First, as already discussed in §3.1, disfluencies involve monitoring and planning by the speaker on various levels. If this is the case, why do speakers make what Clark and Wasow (1998) call apparently premature commitments and start a unit by producing a preposed function

²⁷ More precisely, Levelt and colleagues (1999:5) conceive of the process as follows:

An essential part of the theory, then, is its account of the syllabification process. We have modeled this process by assuming that a morpheme's segments or phonemes become *simultaneously* available, but with labeled links indicating their correct ordering ... The word's metrical template may stay as it is, or be modified in the context. In the generation of *escorting* (or *escort us*, for that matter), the 'spelled out' metrical templates for <escort>, $\sigma\sigma'$, and for <ing> (or <us>), σ , will merge to form the trisyllabic template $\sigma\sigma'\sigma$. The spelled-out segments are successively inserted into the current metrical template, forming phonological syllables 'on the fly': *e-scor-ting* (or *e-scor-tus*). This process follows quite universal rules of syllabification (such as maximization of onset and sonority gradation; see below) as well as language-specific rules.

word, only to halt further production immediately thereafter? Why don't they just wait until they are ready to produce the intended unit without a major disruption?

Second, the asymmetry in production outcomes between preposed and postposed function words does not necessarily explain the further grammaticization of this asymmetry that is manifest in the suffixing preference (as per hypothesis 4) and in the phenomenon of ditropic clitics discussed in §3.2. In order to explain the latter as a direct consequence of the asymmetric production outcomes outlined above, one would have to assume that at some point in the history of the languages showing ditropic clitics, speakers hesitated so frequently after (some classes of) preposed function words that production as part of the preceding prosodic word became the norm rather than the exception. This assumption seems tenuous at best. Instead, one would expect the grammaticization of the 'wrong' alignment to be somehow also functionally motivated. What, then, would be a motivation to systematically phrase preposed function words with the preceding content word?

That there may indeed be a function involved is suggested by experimental evidence reported in Howell & Sackin 2001. They set up a task that involved punishing silent pauses in spontaneous speech, forcing speakers to go on as long as possible without pausing silently. The preferred filling device for pauses of all kinds were repetitions of preposed function words.²⁸ This shows two things. First, it provides additional evidence for the scenario outlined above in that it shows that preposed function words are always ready to be articulated in the syntactic slots they belong in and can be repeated there regardless of the content word following. Second, it suggests that the production of a function word before a prosodic discontinuity in the case of disfluencies has the function of avoiding complete silence and signaling that the speaker has not yet finished and intends to go on.

Clark and Wasow (1998) use the notion of making a PRELIMINARY, rather than just a haphazardly premature, COMMITMENT to designate the function of producing function words before the suspension of speech in disfluencies. That is, by uttering a function word, speakers commit to the further production of speech, and more specifically to 'producing one or more constituents—and to meaning something by them' (Clark & Wasow 1998:237). They do this, according to Clark and Wasow (1998:238; see also p. 208), because

they are pressed by a temporal imperative: They must justify any excessive time they take in speaking (Clark, 1996; Goffman, 1981). If they delay too long, they may be heard as opting out, as confused or distracted, as uncertain about what they want to say, or as having nothing immediately to contribute. They can forestall these attributions by producing the first word of the next constituent (even if prematurely) to show that they are engaged in planning the constituent.

In other words, the preliminary commitment to the further production of speech is of use in achieving interactional goals. This function of uttering a function word prematurely can be made more precise in the turn-taking framework developed by conversation analysts. In their classic article, Sacks, Schegloff, and Jefferson (1974) start with the observation that turn-taking in everyday conversations—at least in white middle-

²⁸ I am not aware of similar experiments involving postposed function words. Fox and colleagues (1996:205) note that Japanese speakers, unlike English speakers, do not appear to use repetitions as a strategy to delay content-word production in conversational speech, attributing this difference to a difference in grammatical structure (positioning the function word before or after the content word). Hayashi (2003a) provides further details on the main strategies used by Japanese speakers when encountering problems in lexical access, which do not include function-word repetitions either. Still, it would be interesting to see whether these findings can be replicated in an experimental setting.

class American English—is characterized by a high degree of smoothness in that there is very little overlap, but a high degree of latching (i.e. the absence of any discernible silence) between the consecutive turns of coparticipants in an interaction. They conclude from this observation that participants in a conversational interaction have to be able to PROJECT the end of a currently unfolding turn, which they call the ‘transition-relevance place’. They remain somewhat vague with regard to which cues hearers use to project the end of a current speaker’s turn, and to this day there is no simple and clear-cut inventory of means of indicating transition-relevance place.

Nevertheless, it is widely agreed that both syntax and prosody play an important role in turn projection. As Auer (1996) argues, syntactic and prosodic cues play different and at times complementary roles in turn projection. That is, at the end of a syntactic unit the prosody can signal incompleteness: the intention of the speaker to further expand the present turn. And vice versa, at the end of a prosodic unit syntactic incompleteness also indicates the intention of the speaker to go on. In this view, the production of a preposed function word before a prosodic discontinuity is an efficient floor-keeping strategy, clearly indicating that the speaker wants to go on. Regularly including in the preceding prosodic word a function word that syntactically forms a constituent with the following content word can then be seen as a grammaticization of such a strategy.

But aren’t postposed function words also conversationally projective and thus equally well suited to become floor-holding devices? And if so, would that not also be a factor hindering fusion with the lexical host? There is no doubt about the fact that postposed function words also have interactional import, as widely documented in the literature on Japanese and Korean conversation. Thus, for example, Tanaka (1999) ‘shows that Japanese participants orient to postpositional particles as a major resource for projecting unfolding turn shapes and organizing turn-taking in Japanese talk-in-interaction’ (Hayashi 2001:338). But there are two important differences. The first one has already been discussed above: the interactional potential of postposed function words can be discharged WITHOUT affecting their prosodic constituency.²⁹ For example, uttering the accusative postposition *o* in the Japanese example in 15 above signals in most contexts that there is at least one more constituent to come that will provide the governor for the phrase thus marked. But this interactional function can be deployed without breaking up the prosodic unit between the clitic function word and its lexical host. Postposed function words almost always form a prosodic unit with their lexical hosts, regardless of the interactional context and momentary suspensions of fluent speech production.

The second difference pertains to the fact that postposed function words generally project to a higher-level constituent (typically, the unfolding clause), never to the lower-level constituent that includes its direct complement. For example, the use of a postposed accusative marker may, in a given context, conversationally project that the governing verb is still to come. But it does not conversationally project the postpositional phrase itself, because this has already been completed by the postpositional marker. Prepositional markers, by contrast, may project both the lower-level phrase they are a part of, and the larger unit (typically a clause) that includes this phrase.

If this second difference is in fact of relevance in the current context, it suggests that we need a more precise delimitation of the type of conversational projectivity at work here. This would allow us to circumscribe precisely the set of function words that are least likely to become further grammaticized as affixes, but that are particularly prone to become ditropic clitics.

²⁹ An exception are the postposition-initiated utterances mentioned in n. 11 above.

It is not possible to discuss this problem in detail within the confines of this article. The remainder of this section merely provides the outline of a possible solution. The basic idea is that the relevant set of function words has to (i) be preposed and (ii) project a specific construction with a specific target that signals the possible syntactic completion of the unit thus projected. Such constructions could be called **TARGET-SPECIFIC CONSTRUCTIONS**. Table 10 lists some English constructions that are target-specific in the intended sense.

MARKER	TARGET
ARTICLE	NOUN (e.g. the bicycle)
AUXILIARY	NONFINITE VERB (e.g. might be going, was done)
PREPOSITION	NOUN (e.g. by chance, of the city, to John)

TABLE 10. Target-specific constructions in English.

What, then, are constructions that are not target-specific in the intended sense? As is immediately obvious, Table 10 includes only constructions below the main clause level. Main clauses (and sentences) are NOT target-specific constructions, because they are projected in a much less rigorous way. As illustrated shortly, there are neither grammatical markers unambiguously signaling the beginning of a clause nor are there target elements that function as typical completion points for clauses. That is, use of an article unambiguously signals the beginning of a noun phrase, which is complete only once the head noun has been uttered. If this does not happen, the unit initiated by the article recognizably remains a fragment, as in example 30.

- (30) and then **the** [.2] and then three boys walk by, (Chafe 1980:319, story 20)

There is no equivalent marker for clause-level units in English. Conjunctions, for example, may initiate a range of constructions, not all of which involve a verbal predicate. Hence they cannot be said to project a verbal predicate in the same way as a definite article projects a noun. Compare examples 31 and 32. The units initiated by the conjunction *but* here are not fragments, although they do not involve a verbal predicate.

- (31) I wonder how the hell he's going carry this basket. (Chafe 1980:303, story 2)

[.3] **But** okay.

- (32) [1.9 [.7] **But** u—h [.35]] still. (Chafe 1980:303, story 2)

/It/ comes back to the old man,

Similarly, while the first NP in an English clause-level unit functions as subject most of the time, this is not necessarily so. For example, initial NPs may also have a topic function and thus not be in direct construction with the following predicate.

- (33) [.6] And **the other little boys** one boy has a little .. bat, (Chafe 1980:303, story 2)

In the following example, the syntactic function of the NP *young boys* is hard to tell. Notably, it is impossible to predict what will happen after it. Still, it is not a fragment, unlike an article or a preposition not followed by a noun.

- (34) .. the film .. flashes to [2.0 [1.45] u—h [.1]] there are three little boys.

[.95] Also about the same age not really necessarily,

one real little,

one [.55] medium sized,

and one about [.25] my height five two?

[.7] **Young boys.**

[.2] Ten,

eleven years old.

[1.35] And they .. help.

(Chafe 1980:312, story 12)

Note that target-specificity does not mean that it is possible to predict precisely how a target-specific construction will unfold after it has been initiated through use of its marker. There may be several elements intervening between marker (*the*) and target (*boys*) as in *the other little boys*, and the construction may be expanded beyond the target (*girl*) as in *a girl on a bicycle*. And, as seen in 30, the construction may be abandoned altogether. But, and this is perhaps the strongest diagnostic for the difference, abandoning a target-specific construction results in a fragment, while non-target-specific constructions may be changed midway without anyone noticing. Thus, for example, the omission of *young boys* or *ten, eleven years old* from example 34 would not be reconstructible. Similarly, continuing *the other little boys* in 33 with *help him* would not show any traces.

Furthermore, to make this distinction does not mean that clauses and other non-target-specific constructions are not also conversationally projectable. But conversational projectability in the case of non-target-specific constructions depends much more strongly on nongrammatical contextual cues (including prosody and pragmatics). For target-specific constructions, conversational projectability is part of their grammatical make-up.

Importantly, not all preposed function words initiate target-specific constructions. One major example are personal pronouns. Of course, these are also often used as preliminary commitments in the sense of Clark & Wasow 1998 and thus frequently occur before hesitation pauses. However, similar to the lexical NP examples above, pronominal NPs do not project a construction with a specific completion target, as examples 35 and 36 illustrate. Note that these examples do not involve false starts, which would typically end in a glottal stop. Instead, the pronouns here are functionally equivalent to the NP *the other little boys* in 33 above, marking a topic shift.

- (35) [.7] one of them .. whistles back to the guy on the bicycle,
 ‘Here’s you hat,’ [sic]
 [.35] or **he** [.4] I don’t know,
 and he goes and takes it, (Chafe 1980:308, story 9)
- (36) [.5] and apparently **he** [.9] I think by the breeze, (Chafe 1980:309, story 10)
 .. his hat sort of gets [.7] blown off his head so he sort of turns around,

The proposed distinction between initiating a target-specific construction and being conversationally projective in a more general, not construction-specific sense may then also provide a venue to explain the differences between preposed function words with regard to the suffixing preference noted in §1 above. The prediction is that the more clearly a given preposed function word projects a target-specific construction, the less likely it is to be further grammaticized as an affix.

This prediction is based on the following reasoning. A preposed function word that initiates a target-specific construction is a particularly efficient floor-keeping device since it is a clear signal of the intention of the current speaker to go on. That is, it may be conversationally useful to utter such a function word even when the planning and computation of subsequent segments has not yet been finished. The fact that they can be used in this way at all is due to their high frequency and the incremental left-to-right nature of speech production. As a consequence, preposed function words can be, and are, uttered prematurely whenever the preceding segment is done. This in turn means that they are frequently separated by a prosodic boundary from their lexical hosts. The prosodic boundary may consist in a disfluency. But it may also be more grammaticized and involve the inclusion of the preposed function word in the preceding phonological

word, resulting in a ditropic clitic. Both types of boundaries prevent the further grammaticization of (clitic) function words as affixes on their lexical hosts.

The potential for premature production is common to all preposed function words. But this potential is regularly discharged only in the case of function words initiating target-specific constructions. Function words not initiating target-specific constructions are less strongly conversationally projective. Hence it is conversationally less useful to utter them prematurely. If this hypothesis and the scenario just sketched are correct, function words that do NOT initiate target-specific constructions should be equally likely to further grammaticize to affixes regardless of their position before or after their hosts. And this is indeed what we find: the affixation rates for preposed case markers (which develop from adpositions), preposed tense-aspect-mood markers, and preposed plural markers are very low (see Table 3 above), but they are about equal for personal pronouns (see §1).³⁰

5. CONCLUSION. This article puts forth the hypothesis that the well-known asymmetry in the rate of affixation of grammatical elements in the languages of the world—postposed markers are much more likely to be affixes than preposed ones—is related to other asymmetries between preposed and postposed function words with regard to their prosodic chunking. Specifically, it is shown that preposed function words are more often and more systematically separated from their lexical hosts by a prosodic boundary. This is the case both when disfluencies in spontaneous speech occur (§3.1) and when a misalignment between syntactic and prosodic boundaries occurs (ditropic clitics are always preposed function words, never postposed ones; §3.2). The prosodic boundary, according to the hypothesis, is a major factor in blocking the fusion of a preposed function word and its lexical host.

The regular and frequent occurrence of a prosodic boundary after preposed function words has a natural basis in the mechanics of speech production. Function words tend to be high-frequency forms that are often also highly predictable in a given context. They are thus always highly activated and ready to go into production, whenever the incremental planning and execution of an utterance has reached the point where they are to be pronounced. This, in turn, provides the possibility of uttering a function word even when the speaker is not yet fully done with computing the next unit in the overall utterance plan.

That preposed function words are in fact uttered prematurely, as it were, is due to the fact that by doing this, speakers may achieve interactional goals. Most importantly, they signal their intention of continuing and thus defending their right to hold the floor as the current speakers.

Not all preposed function words are equally well suited for achieving these interactional goals. In §4 it is suggested that there is one type of preposed function word that is particularly well suited for the purpose. These are function words that initiate target-specific constructions, that is, constructions that have a clearly defined syntactic endpoint and thus remain incomplete when this endpoint is not reached. For example, use of a definite article projects the subsequent occurrence of a noun, and the construction it initiates is incomplete as long as the noun has not been provided.

It is hypothesized that function words initiating target-specific constructions are particularly resistant to further grammaticization as affixes because they are particularly

³⁰ Unfortunately, we do not have the relevant data for articles. Maps 37 and 38 for definite and indefinite articles, respectively, in Dryer & Haspelmath 2011 distinguish between definite/indefinite ‘words’ and ‘affixes’, but do not provide information on the position of these definiteness markers relative to their hosts.

useful in projecting an upcoming transition-relevance place. Postposed function words do not generally initiate target-specific constructions (but rather conclude them) and thus, according to this hypothesis, are more likely to grammaticize, since they are also not generally separated by prosodic boundaries from their lexical hosts. But there are also preposed function words that do not initiate target-specific constructions, for example, personal pronouns; these are also predicted to be more amenable to fusion than the ones initiating such constructions.

The hypothesis proposed here further develops the earlier proposals for explaining the suffixing preference reviewed in §2. From the approach proposed by Cutler, Hawkins, and Gilligan (1985) and further refined by Hall (1992), it takes on the idea that the suffixing preference is related to aspects of language processing. But while these authors try to link it to aspects of language COMPREHENSION, specifically word comprehension (stems are processed earlier than affixes), the current scenario links it to aspects of language PRODUCTION (the high activation of function words allows them to go into production as soon as the preceding element in the incremental left-to-right processing of an utterance has been uttered).

The major advantage of this scenario is that it does not require postulating a fundamental difference between preposed and postposed function words in terms of processing. The processing account outlined in §4 holds for both types of function words. The different outcomes with regard to prosodic discontinuities are simply due to the different positions with regard to their hosts, as illustrated in Table 11.

PREPOSED	X	function word	prosodic boundary	complement (= possible projection target)
POSTPOSED	complement	function word	prosodic boundary	X

TABLE 11. Different implications of prosodic discontinuities after pre- and postposed function words.

Thus, unlike Hall 1992, the current account does not have to assume a rather speculative ‘flirting’ process, where preposed function words continually are on their way to fusing with the stem but never actually fuse because of the preferred computational order in word recognition (stem before affixes).

From the work of Bybee and colleagues (1990), the current approach adopts the fundamental distinction between a general postposing tendency (grammatical elements are more frequently found after their hosts than before them) and the suffixing preference proper (postposed grammatical elements are more frequently fused with their hosts than preposed ones). The current approach has nothing to contribute to explaining the postposing tendency, which in all likelihood reflects general word-order tendencies in the languages of the world (SOV order is much more frequent than VSO/VOS). Instead, the hypothesis developed here becomes relevant only when the function words are already in the clitic stage of the grammaticization cline in 1 and hence ‘in place’ (before or after their hosts).

Previous work, by Hall (1992:162) among others, has noted that in order for fusion to occur between a grammatical element and its host, the two items have to occur adjacent to each other with sufficient frequency. In line with this work, the current approach has assumed that no major differences exist between preposed and postposed function words with regard to adjacency requirements. It should be noted, however, that this issue has not been properly investigated thus far. Thus, it is possible that there are in fact statistically significant differences with regard to the adjacency requirements for

preposed and postposed function words, and that these differences also contribute to the suffixing preference.

Still, as Bybee and colleagues (1990:29) remark, exclusive adjacency is not enough for fusion to occur. Instead, a grammatical element and its host have to regularly occur within the same prosodic unit for this to happen. To date, this requirement has not been further investigated. The hypothesis in 4—that preposed function words are more often separated from their lexical hosts by a prosodic boundary than postposed ones—is thus the core of the innovative approach presented here. This, in turn, is directly linked to the further hypotheses that on a deeper level, the suffixing preference is related to the mechanics of speech production and conversational floor-keeping strategies. Looking closely at the prosody uncovered the asymmetry in ditropic clitics, which phonologically always attach to the left (the preceding element). This asymmetry had never previously been explicitly noted nor had it been linked to the suffixing asymmetry.

Inasmuch as function words and their hosts tend to occur within the same prosodic unit at clause boundaries, the current proposal is congruent with the proposal by Bybee and colleagues (1990:34) that the rate of affixation is particularly high at clause boundaries. However, the claim here is that it is not clause boundaries *per se* that promote the further grammaticization of clitics to affixes. Instead, what is relevant is the fact that clause boundaries are regularly also prosodic boundaries.

According to the scenario proposed in this article, prosodic boundaries are ONE among a number of factors relevant for the suffixing preference. Semantic relevance is a key factor as well, as cogently argued in Bybee et al. 1990:29–34. Function word and lexical host have to form a coherent conceptual unit for fusion to occur. As already noted in n. 26, semantic relevance is needed to explain the fact that ditropic clitics rarely, if ever, develop into suffixes on their phonological hosts, because the semantic bond between them and their morphosyntactic hosts is much closer (articles are much more closely linked to the noun they determine than to a verb that may happen to precede them and that they may lean on enclitically).

The current approach is not (yet) capable of explaining all data pertaining to the suffixing preference. Specifically, the tendency of preposed function words to be more frequently affixes in SOV languages than in SVO languages is not explained. This tendency is apparent in the database of Bybee et al. (see columns 2 and 4 in Table 2). It has not been further discussed in the literature and is still in need of independent corroboration. Note that Bybee and colleagues' own proposal for explaining the suffixing preference—affixation rates are higher at clause boundaries than clause-medially (1990:34)—does not explain this tendency either, because in both instances the preposed function words do occur clause-medially (before V and after S or O).

Unlike previous approaches, however, the current scenario provides ample room for crosslinguistic variation, which is necessary to account for the fact that prefixes occur in the languages of the world and, more specifically, that a few languages are in fact heavily prefixing. It allows for crosslinguistic variation on two levels. On the one hand, it may be possible that speech communities differ with regard to the temporal pressures applied on current speakers and thus have weaker requirements for them to show themselves as intending to go on with an utterance currently under way. On the other hand, and probably more importantly, speech communities may differ as to which devices are conventionalized for signaling the intention to hold the floor. We still do not know very much with regard to either of these dimensions of possibly culture-specific variation, but there is some evidence that such variation does in fact exist. Stivers and colleagues (2009) provide data and discussion on timing differences in turn-taking across ten lan-

guages. The fact mentioned above that in western Austronesian languages there are no conventional pause fillers such as *uh* (superlengthening of grammatical elements being the standard alternative) is a case in point for the second dimension.

REFERENCES

- ANDERSON, STEPHEN R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.
- ANDERSON, STEPHEN R. 2005. *Aspects of the theory of clitics*. Oxford: Oxford University Press.
- AUER, PETER. 1996. On the prosody and syntax of turntaking. *Prosody in conversation: Interactional studies*, ed. by Elizabeth Couper-Kuhlen and Margret Selting, 57–100. Cambridge: Cambridge University Press.
- BEATTIE, GEOFFREY W., and B. L. BUTTERWORTH. 1979. Contextual probability and word frequency as determinants of pauses and speech errors in spontaneous speech. *Language and Speech* 22.201–11.
- BYBEE, JOAN L.; WILLIAM PAGLIUCA; and REVERE D. PERKINS. 1990. On the asymmetries in the affixation of grammatical material. *Studies in typology and diachrony: Papers presented to Joseph H. Greenberg on his 75th birthday*, ed. by William A. Croft, Suzanne Kemmer, and Keith Denning, 1–42. Amsterdam: John Benjamins.
- CHAFE, WALLACE L. (ed.) 1980. *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*. Norwood, NJ: Ablex.
- CLARK, HERBERT H. 1996. *Using language*. Cambridge: Cambridge University Press.
- CLARK, HERBERT H., and JEAN E. FOX TREE. 2002. Using *uh* and *uhm* in spontaneous speaking. *Cognition* 84.73–111.
- CLARK, HERBERT H., and BARBARA C. MALT. 1984. Psychological constraints on language: A commentary on Bresnan and Kaplan and on Givón. *Method and tactics in cognitive science*, ed. by Walter Kintsch, James R. Miller, and Peter G. Polso, 191–214. Hillsdale, NJ: Lawrence Erlbaum.
- CLARK, HERBERT H., and THOMAS WASOW. 1998. Repeating words in spontaneous speech. *Cognitive Psychology* 37.201–42.
- CROFT, WILLIAM. 1995. Intonation units and grammatical structure. *Linguistics* 33.839–82.
- CUTLER, ANN; JOHN HAWKINS; and GARY GILLIGAN. 1985. The suffixing preference: A processing explanation. *Linguistics* 23.723–58.
- CYSOUW, MICHAEL. 2005. Morphology in the wrong place: A survey of preposed enclitics. *Morphology and its demarcations*, ed. by Wolfgang U. Dressler, 17–37. Amsterdam: John Benjamins.
- CYSOUW, MICHAEL. 2009. The asymmetry of affixation. *Snippets* (Special issue in honor of Manfred Krifka, ed. by Sigrid Beck and Hans-Martin Gärtner) 20.10–14. Online: <http://www.ledonline.it/snippets/>.
- DIEZ, FRIEDRICH. 1838. *Grammatik der romanischen Sprachen*, zweiter Theil. Bonn: Weber.
- DRYER, MATTHEW S. 2011a. Prefixing vs. suffixing in inflectional morphology. In Dryer & Haspelmath, Ch. 26. Online: <http://wals.info/chapter/26>.
- DRYER, MATTHEW S. 2011b. Coding of nominal plurality. In Dryer & Haspelmath, Ch. 33. Online: <http://wals.info/chapter/33>.
- DRYER, MATTHEW S. 2011c. Position of case affixes. In Dryer & Haspelmath, Ch. 51. Online: <http://wals.info/chapter/51>.
- DRYER, MATTHEW S. 2011d. Position of pronominal possessive affixes. In Dryer & Haspelmath, Ch. 57. Online: <http://wals.info/chapter/57>.
- DRYER, MATTHEW S. 2011e. Position of tense-aspect affixes. In Dryer & Haspelmath, Ch. 69. Online: <http://wals.info/chapter/69>.
- DRYER, MATTHEW S., and MARTIN HASPELMATH (eds.) 2011. *The world atlas of language structures online*. Munich: Max Planck Digital Library.
- DU BOIS, JOHN W.; WALLACE L. CHAFE; CHARLES MEYER; and SANDRA A. THOMPSON. 2000. *Santa Barbara corpus of spoken American English, part 1*. Philadelphia: Linguistic Data Consortium.
- DU BOIS, JOHN W.; WALLACE L. CHAFE; CHARLES MEYER; SANDRA A. THOMPSON; and NII MARTEY. 2003. *Santa Barbara corpus of spoken American English, part 2*. Philadelphia: Linguistic Data Consortium.

- DU BOIS, JOHN W., and ROBERT ENGLEBRETSON. 2004. *Santa Barbara corpus of spoken American English, part 3*. Philadelphia: Linguistic Data Consortium.
- DU BOIS, JOHN W., and ROBERT ENGLEBRETSON. 2005. *Santa Barbara corpus of spoken American English, part 4*. Philadelphia: Linguistic Data Consortium.
- DU BOIS, JOHN W.; STEPHAN SCHUETZE-COBURN; SUSANNA CUMMING; and DANAE PAOLINO. 1993. Outline of discourse transcription. *Talking data: Transcription and coding in discourse research*, ed. by Jane Edwards and Martin D. Lampert, 45–89. Hillsdale, NJ: Lawrence Erlbaum.
- FORTESCUE, MICHAEL. 2006. Drift and the grammaticalization divide between Northern and Southern Wakashan. *International Journal of American Linguistics* 72.295–324.
- FOX, BARBARA A.; MAKOTO HAYASHI; and ROBERT JASPERSON. 1996. Resources and repair: A cross-linguistic study of syntax and repair. *Interaction and grammar*, ed. by Elinor Ochs, Emanuel A. Schegloff, and Sandra A. Thompson, 185–237. Cambridge: Cambridge University Press.
- GOFFMAN, ERVING. 1981. Radio talk. *Forms of talk*, 197–327. Philadelphia: University of Pennsylvania Press.
- GOLDMAN EISLER, FRIEDA. 1968. *Psycholinguistics: Experiments in spontaneous speech*. New York: Academic Press.
- GREENBERG, JOSEPH H. 1957. Order of affixing: A study in general linguistics. *Essays in linguistics*, 86–97. Chicago: University of Chicago Press.
- GUSSENHOVEN, CARLOS, and HAIKE JACOBS. 1998. *Understanding phonology*. London: Arnold.
- HALL, CHRISTOPHER J. 1988. Integrating diachronic and processing principles in explaining the suffixing preference. *Explaining language universals*, ed. by John A. Hawkins, 321–49. Oxford: Oxford University Press.
- HALL, CHRISTOPHER J. 1992. *Morphology and mind: A unified approach to explanation in linguistics*. London: Routledge.
- HAWKINS, JOHN A., and ANNE CUTLER. 1988. Psycholinguistic factors in morphological asymmetry. *Explaining language universals*, ed. by John A. Hawkins, 280–317. Oxford: Oxford University Press.
- HAWKINS, JOHN A., and GARY GILLIGAN. 1988. Prefixing and suffixing universals in relation to basic word order. *Lingua* 74.219–59.
- HAYASHI, MAKOTO. 2001. Postposition-initiated utterances in Japanese conversation: An interactional account of a grammatical practice. *Studies in interactional linguistics*, ed. by Margret Selting and Elizabeth Couper-Kuhlen, 317–43. Amsterdam: John Benjamins.
- HAYASHI, MAKOTO. 2003a. Language and the body as resources for collaborative action: A study of word searches in Japanese conversation. *Research on Language and Social Interaction* 36.109–41.
- HAYASHI, MAKOTO. 2003b. *Practices in joint utterance construction in Japanese conversation*. Amsterdam: John Benjamins.
- HOPPER, PAUL J., and ELIZABETH CLOSS TRAUGOTT. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.
- HOWELL, PETER, and STEVIE SACKIN. 2001. Function word repetitions emerge when speakers are operantly conditioned to reduce frequency of silent pauses. *Journal of Psycholinguistic Research* 30.457–74.
- IWASAKI, SHOICHI. 1993. *Subjectivity in grammar and discourse*. Amsterdam: John Benjamins.
- KLAVANS, JUDITH. 1985. The independence of syntax and phonology in cliticization. *Language* 61.95–120.
- LEHMANN, CHRISTIAN. 1995. *Thoughts on grammaticalization*. Munich: LINCOM Europa.
- LEVELT, WILLEM J. M. 1983. Monitoring and self-repair in speech. *Cognition* 14.41–104.
- LEVELT, WILLEM J. M. 1989. *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- LEVELT, WILLEM J. M. 1992. Accessing words in speech production: Stages, processes and representations. *Cognition* 42.1–22.
- LEVELT, WILLEM J. M.; ARDI ROELOFS; and ANTJE S. MEYER. 1999. A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22.1–38.
- LOCAL, JOHN, and JOHN KELLY. 1986. Projection and ‘silences’: Notes on phonetic and conversational structure. *Human Studies* 9.185–204.

- MARSLEN-WILSON, W. D. 1987. Functional parallelism in spoken word recognition. *Cognition* 25.71–102.
- MATSUMOTO, KAZUKO. 2000. Japanese intonation units and syntactic structure. *Studies in Language* 24.515–64.
- MITHUN, MARIANNE. 2003. Why prefixes? *Acta Linguistica Hungarica* 50.155–85.
- MORITA, EMI. 2005. *Negotiation of contingent talk: The Japanese interactional particles ne and sa*. Amsterdam: John Benjamins.
- NESPOR, MARINA, and IRENE VOGEL. 1986. *Prosodic phonology*. Dordrecht: Foris.
- REID, LAWRENCE A. 1970. *Central Bontoc: Sentence, paragraph and discourse*. Norman: University of Oklahoma and Summer Institute of Linguistics.
- SACKS, HARVEY; EMANUEL SCHEGLOFF; and GAIL JEFFERSON. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50.696–735.
- SAPIR, EDWARD. 1921. *Language: An introduction to the study of speech*. New York: Harcourt, Brace.
- SELKIRK, ELISABETH. 1995. The prosodic structure of function words. *Papers in optimality theory* (University of Massachusetts occasional papers in linguistics 18), ed. by Jill Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk, 439–70. Amherst: GLSA Publications.
- SIMPSON, ADRIAN P. 2006. Hesitation phenomena and pauses. *Encyclopedia of language and linguistics*, 2nd edn., ed. by Keith Brown, vol. 5, 284–88. Amsterdam: Elsevier.
- SPENCER, ANDREW. 1991. *Morphological theory*. Oxford: Blackwell.
- SPENCER, ANDREW, and ANA R. LUÍS. 2012. *Clitics: An introduction*. Cambridge: Cambridge University Press.
- STIVERS, TANJA; N. J. ENFIELD; PENELOPE BROWN; CHRISTINA ENGLER; MAKOTO HAYASHI; TRINE HEINEMANN; GERTIE HOYMAN; FEDERICO ROSSANO; JAN PETER DE RUITER; KYUNG-EUN YOON; and STEPHEN C. LEVINSON. 2009. Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences* 106.10587–92. Online: <http://www.pnas.org/content/106/26/10587.full>.
- STREECK, JÜRGEN. 1996. A little Ilokano grammar as it appears in interaction. *Journal of Pragmatics* 26.189–213.
- TANAKA, HIROKO. 1999. *Turn-taking in Japanese conversation: A study in grammar and interaction*. Amsterdam: John Benjamins.
- ZWICKY, ARNOLD M., and GEOFFREY K. PULLUM. 1983. Cliticization vs. inflection: English *n't*. *Language* 59.502–13.

Allgemeine Sprachwissenschaft
 Institut für Linguistik
 Universität zu Köln
 50923 Köln, Germany
sprachwissenschaft@uni-koeln.de

[Received 18 December 2011;
 revision invited 28 January 2013;
 revision received 20 July 2013;
 accepted 31 January 2014]