J. Linguistics **42** (2006), 289–315. © 2006 Cambridge University Press
doi:10.1017/S0022226706003902 Printed in the United Kingdom

Trapped morphology[1]

ALICE C. HARRIS

SUNY Stony Brook, Center for Advanced Study

JAN TERJE FAARLUND

Centre for Advanced Study, University of Oslo

(Received 25 July 2005; revised 12 December 2005)

We argue that there is a diachronic process, distinct from phonological erosion, that results in the loss of inflectional morphology that is trapped when a clitic attaches to a host, becoming an affix. This is supported with attested examples from Mainland Scandinavian, Georgian, Spanish, and Greek, as well as shallow, well-accepted reconstructions from Slavic and Georgian. It is further supported by new reconstructions from Zoque (Mixe-Zoquean) and Andi (Northeast Caucasian). For example, in Old Norse the postposed article is a clitic, and there is a case ending between the noun stem and the article: *hest-s=in-s* 'the horse (gen)'. The first *s* is trapped morphology, and it is subsequently lost: *hest-en-s*. Similarly, in pre-Georgian, the postposed article traps the ergative case marker, *-*n*: **k'ac-n=ma-n* 'the man (erg)'; it is subsequently lost: *k'ac-man*. We argue that the loss of trapped morphology is not sound change or another phonological process, but a morphological process.

1. Introduction

In recent years there has been considerable interest in the process of grammaticalization, which includes the process of words becoming clitics and clitics becoming affixes. In this paper, we argue that even with this considerable attention, one aspect of the progression of words to affixes has been insufficiently studied. We suggest that there is a morphological process that accounts for the loss of morphemes trapped between a word and a clitic. We argue that this process is widespread, but does not apply invariably when morphemes are trapped.

[1] The research reported here was supported in part by the National Science Foundation under Grant Number BCS-0215523 (ACH) and by the Centre for Advanced Study in Oslo in 2005 (both authors). We are grateful to our colleagues at the Centre for Advanced Study for their comments, especially to Henning Andersen for help with the Slavic data. We also appreciate comments and challenges from members of the audience at the workshop Weak Words: Their Origin and Progress, held in Konstanz in April 2005, where an earlier version of this paper was presented. We would also like to thank Brian Joseph for discussion of the Greek data.

We consider inflection to be TRAPPED if it occurs between a stem and a clitic, as schematized in schema 1.

Prefix-	HOST	-Suffix	Prefix-	CLITIC	-Suffix
Irrelevant		Trapped Morphology			Irrelevant

Schema 1
Trapped morphology

Although schema 1 displays only enclitics, we assume that this is a general process that applies with proclitics as well as with enclitics.

In a simple case, if a verb has agreement affixes, the auxiliary often has the same agreement affixes. If the verb and auxiliary become a single word, we often find one of the following two scenarios:[2]

SUBJ-	VERB	SUBJ-	AUX

↘

SUBJ-	VERB		-AUX

Schema 2
Morpheme loss with agreement prefixes

VERB	-SUBJ	AUX	-SUBJ

↘

VERB		-AUX	-SUBJ

Schema 3
Morpheme loss with agreement suffixes

Similarly, in a simple case, if a noun has case suffixes, pronouns in the same language will frequently have similar case suffixes. Pronouns are often

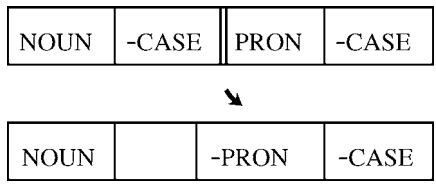
[2] The following abbreviations are used in glossing examples:

3SG = 3rd person singular	DEF = definite	NOM = nominative
3PL = 3rd person plural	DET = determiner	PL = plural
3PS = 3rd person possessive	ERG = ergative	PRON = pronoun
ABSL = absolutive	FEM = feminine	PS = possessive
ACC = accusative	GEN = genitive	SG = singular
AUX = auxiliary	LOC = locative	SUBJ = subject
DAT = dative	MASC = masculine	

In the text, the following are also used: CM = class marker, LPSl = Late Proto-Slavic, OCS = Old Church Slavonic.

https://doi.org/10.1017/S0022226706003902 Published online by Cambridge University Press

grammaticalized as affixes on nouns (Greenberg 1978), and this may lead to the following situation:



Schema 4
Morpheme loss with case suffixes

We argue that this loss is not sound change, erosion, or any other phonological process, but a morphological process. We suggest that the erosion that accompanies grammaticalization is poorly defined and is much too broad. A first step in better understanding the reduction that accompanies grammaticalization is to distinguish the loss of trapped morphemes, which can be rather tightly circumscribed, from erosion, which still appears diffuse. We suggest that loss of trapped morphology differs from erosion in at least three respects. First, the target of the former is typically entire morphemes, while that of the latter is not constrained in this way. Second, the morphological process is limited to trapped inflection, as defined above, while erosion typically targets the clitic and its affixes. Third, erosion is driven by prosody, phonotactics, or general sound change, while the loss of trapped morphemes is unrelated to these processes.

It is true that in some languages phonological differences among morphemes may determine which are lost and which remain (see sections 2.2, 4, and 6). On the basis of the generality of the process, the fact that it targets whole morphemes, and the fact that it is not part of any phonological process that has been clearly identified, we argue that this is a specifically morphological process, which involves loss of inflectional morphology.

One might hypothesize that identity avoidance (Yip 1998 and sources cited there) is a factor in this, given that most of our examples involve loss of a morpheme identical to some other morpheme in the same, newly formed word. However, some examples do not involve either phonological or morphological identity between morphemes, at least in a straightforward way. Thus, in the Zoque example below (section 10), the lost morphemes are not identical in any sense to other morphemes in the word. In the Old Georgian example (section 5), the lost *v-* is identical to another *v-* in the same word, both phonologically and morphologically. However, the *v-* that was not lost was inserted in historical times. Thus, another, simpler method of avoiding identity in this example would have been to do nothing; without insertion of *v-* there would have been no identity, and therefore no trigger for loss, according to the identity avoidance hypothesis.

https://doi.org/10.1017/S0022226706003902 Published online by Cambridge University Press

We believe that the primary conditioning factor in these examples was not identity avoidance but establishing optimal morpheme order (see also Harris 2004). The loss of contextual inflectional morphemes 'inside' (closer to the root than) morphemes that are in the process of being reanalyzed in various derivational or inherent inflectional functions (e.g. tense, mood) avoids dispreferred morpheme order while leaving other inflectional morphemes intact and *in situ* in the preferred 'outside' position.[3]

In order to establish the generality of this change and examine its nature, we adduce examples from a variety of languages. We begin with seven attested examples and two shallow reconstructions that represent standard analyses. These include three examples of nouns or adjectives followed by a determiner, then four examples of verbs with an auxiliary or a reflexive marker. The examples include the loss of the Scandinavian case and plural marker trapped between a noun and article (section 2); the ergative case marker trapped in Georgian between the noun and article (section 3); Slavic long-form adjectives, in which adjective cases are lost when trapped between the adjective and the pronoun (section 4); Georgian person marking trapped between a verb and auxiliary (section 5); Greek person marking similarly trapped (section 6); Spanish person-number marking trapped in a similar way (section 7); and Scandinavian person-number markers trapped between a verb and reflexive (section 8). In order to show that the process is not limited to the Indo-European and Kartvelian families, we include reconstructed examples from Andi (Northeast Caucasian family, section 9) and Zoque (Mixe-Zoquean, section 10). We end with a discussion in section 11 of other possible outcomes, and a conclusion in section 12.

2. Scandinavian definite article

2.1 *Trapped case*

In Scandinavian languages the definite article occurs as an ending on the noun. This ending was affixed to the inflected form of the noun in Old Scandinavian (Faarlund 2004: 39, 57f.). The indefinite and definite inflections of a strong masculine noun, 'horse', in the singular are shown in table 1.

The definite noun thus had redundant case marking. (The final *-n* in the definite nominative of masculine nouns, as in *hestrinn*, is the result of assimilation of the second *r* to the preceding *n* of the cliticized determiner: **nr* > *nn*.)

In Mainland Scandinavian all the case endings except the genitive -s are gone, but the *s* now does not intervene between the nominal stem and

[3] For the distinction between contextual and inherent inflectional morphology, see Booij (1994, 1996).

	INDEFINITE DECLENSION		DEFINITE DECLENSION			
	Noun	-Case	Noun	-Case	Article	-Case
NOM	hest	-r	hest	-r	in	-n
ACC	hest		hest		in	-n
DAT	hest	-i	hest	-i	n	-um
GEN	hest	-s	hest	-s	in	-s

Table 1
Indefinite and definite inflection in Old Scandinavian

the definite article. It can only follow the article (except in a few fossilized forms).

(1) hest-en-s 'of the horse'

We claim that the case endings of the noun in the definite inflection, those in the shaded column in table 1, have been systematically lost.

Segments in a word may of course be lost for reasons other than being 'trapped'. As is well known, case inflection in Mainland Scandinavian was lost as a category, not only in trapped positions, but generally. One may therefore ask whether the loss of trapped case morphemes was just a result of the general loss of case affixes, and this question can be answered through study of the chronology of the changes. The best testing ground is Swedish, because it is the Scandinavian language which is best documented for the relevant period, and because the loss of case marking in Swedish has been thoroughly studied and analyzed by Muriel Norde (Norde 1997).

The genitive suffix has the following allomorphs in Old Swedish:

(2) -s (masc., neuter)
-ar (masc., fem.)
-a (plural, all genders)

(3) hæst-s 'horse'
hæst-a 'horses'
færþ-ar 'journey'

The allomorph *-s* is the only one that still exists. In fact it is the only remnant of the old case system in Mainland Scandinavian still in productive use, and it is now used only as a marker of possession, having lost its other genitive functions. In the position of interest to us, the shaded column of table 1, the suffix *-s* was also lost, but it continues to be used after the article and after the bare stem.

https://doi.org/10.1017/S0022226706003902 Published online by Cambridge University Press

(4) en hest-s 'of a horse'
hest-en-s 'of the horse'

Thus, loss of *-s* in the trapped position, shaded in table 1, cannot be due to the general loss of case, since this one case was in fact preserved.

It could now be argued that the *-s* is no longer an affix; as illustrated in (5b), it has become a phrasal clitic in all the Mainland Scandinavian languages, as in English. But the *-s* was lost in the trapped position before the *-s* at the periphery became a phrasal clitic, in Swedish as well as in Danish and Dano-Norwegian. This stage is preserved in the Swedish example (5a); this is now archaic and is characterized as 'formal written language' by Teleman et al. (1999: 131). The current and less formal version in all the Modern Scandinavian languages would be (5b), with the *-s* attached to the final word of the phrase, as in English.

(5) (a) institution-en-s för slaviska språk prefekt
department-DEF-GEN of Slavic languages chairperson
'the head of the Slavic Department'
(b) institutionen för slaviska språks prefekt

Our claim that loss of trapped *-s* is independent of the development of *-s* into a phrasal clitic is further supported by dialect data. There are still regional dialects of Swedish where the old genitive *-s* cannot function as a phrasal clitic. One such variety is the Lappträsk dialect from Nyland in southern Finland, where the equivalent of (5b) would be ungrammatical; thus there is no phrasal genitive in this dialect (Vangsnes 1998). This dialect still has a genitive in *-s*, which occurs only at the end of the word, 'outside' the definite article. Compare the standard Swedish examples in (4) and the Nyland examples in (6) (from Lundström 1939).

(6) (a) po Gröndal-s oker-n
on Gröndal-GEN field-DEF
'in Mr Gröndal's field'
(b) po grannar-s marč-in
on neighbors-GEN ground-DEF
'on the ground of (some) neighbors'
(c) Hagnäs-härr-n-s kuddo-na
Hagnäs-owner-DEF-GEN cows-DEF
'the Hagnäs farmer's cows'

These examples show clearly that the genitive case marker is retained on bare nouns (6a, b) and as a final suffix after the definite article (6c). This is important because in this dialect this case has not been lost in general, nor has it become a phrasal clitic; rather it has been lost just in the trapped position – between the noun stem and the definite article, as shown in (6c). This provides further evidence that the loss of the trapped *-s* is independent of the development of the peripheral *-s* into a clitic.

Norde (1997) documents the fact that from the Old Swedish period onward there was a gradual loss of trapped inflection. The rate and order of the loss depended mainly on the declensional class. The feminine genitive *-ar* provides another instance of loss of trapped inflection. This loss is documented in Old Swedish, while the case system was still intact. Norde (1997: 111f.) gives the following examples, among others, illustrating the loss of the trapped case, while the external case marker was retained:

(7) (a) *Strong feminine*
færþ-in-nar 'the journey' (GEN) < færþ-ar + in-nar
(b) *Strong masculine*
præstins 'the priest' (GEN) < præst-s + in-s

The loss of trapped morphology here cannot be attributed to phonological processes. A phonological account of these examples would entail that the relevant phonological rules apply elsewhere in the Scandinavian languages, but this does not seem to be the case. The cliticized forms would constitute a phonotactically acceptable structure even without the loss of morphology. The *-s* genitive was lost in the trapped position, but it was kept after the same consonant (cluster) when it was not followed by the article. A form such as *hestsens* (the modern equivalent of the Old Scandinavian definite, *hestsins*) ought to be phonotactically more acceptable than *hests* (the actual modern indefinite form), since in the former case the second *s* is the onset of the next syllable, whereas in the latter the *s* forms part of a final consonant cluster. Nevertheless, the trapped *s* of *hestsens* was lost.

2.2 *The trapped plural marker*

In Old Norse the plural suffix had the form *-Vr* in the nominative masculine and in the nominative and accusative feminine. This suffix preceded the definite article, and it too was reduced as the definite article became an affix.[4] It is retained at the end of the indefinite noun, but it is lost in the definite form:

(8) hester 'horses' hestene (< hest-er + ne) 'the horses'

The loss illustrated in (8) may appear to be a phonologically conditioned simplification of the cluster *-rn-* in unstressed syllables, but note that the same cluster is not reduced if the *-r* is not part of a plural suffix:

(9) baker 'baker' bakerne (< baker + er + ne) 'the bakers'

[4] It can be argued that the *-r* of the plural was really a case suffix. This may be a correct analysis for Old Norse, but in Modern Norwegian, where the case inflection is now lost, this *-r* is the marker of plural.

In a deverbal noun like *bak-er* the final *-r* of the agentive suffix is not lost. However, the entire plural suffix *-er* is lost between the agentive suffix (*-er*) and the definite article (*-ne*).[5]

We conclude that trapped plural markers in Old Scandinavian, like trapped case markers, were lost morphologically.

3. Loss of trapped ergative case in Georgian

Georgian is a language of the Kartvelian family, and Old Georgian dates from the fourth or fifth century CE. In Old Georgian, definite nouns were immediately followed by a definite article, as illustrated in (10), with the noun *saxl-* 'house'.

(10)

	SINGULAR		COLLECTIVE		PLURAL	
NOM	saxl-i	igi	saxl-eb-i	igi	saxl-ni	igi
ERG	saxl-man	man	saxl-eb-man	man	saxl-ta	mat

The article of Old Georgian is identical to the demonstrative, except that a singular form occurs as the article in the definite nominative plural. The demonstrative occurs in four series of forms: proximal, medial, remote, and neutral.[6] The example in (10) is slightly simplified, in that other forms of the article, with the prefix *a-* or *i-*, might also have been the input to change.

Čikobava (1939), Šaniʒe (1957), Mač'avariani (1960, 1985), and others have shown that at an earlier stage the determiner grammaticalized as nominative and ergative case markers, as illustrated for the singular in (11).

(11)

NOM	saxl-i	< *saxl	igi
ERG	saxl-man	< *saxl-n	man

[5] The indefinite plural of this class of nouns ends in *-er+er* in East Norwegian dialects (*bakerer* 'bakers'), but in the standard language the final *-r* is lost, yielding *bakere* – presumably under Danish influence, or as a phonological process, or both.

[6] The four sets of forms in the singular are shown in (i).

(i)

	PROXIMAL	MEDIAL	REMOTE	NEUTRAL
NOM	ese	ege	(isi)	igi
ERG	aman	magan	(iman)	man
DAT	amas	magas	imas	mas
GEN	amis	magis	imis	mis
INST	amit	magit	imit	mit

While the remote forms in parentheses are attested in Old Georgian only as demonstratives, not as definite articles, the remote dative is not infrequent, and this suggests that remote forms may have been more common in pre-Old Georgian (Imnaišvili 1955). Neutral (perhaps more appropriately considered an alternant of the remote, *ibid.*), proximal, and medial forms could also function as articles. Thus, the specifically proximal forms in (10) are not intended as the only possibility, and we must bear in mind that the deictic element, *a-*, may not have been present in the input to the forms attested. Further, the forms represented in (10) are a mixed paradigm, containing elements of both the neutral (nominative form) and proximal sets (other case forms).

In the nominative, no overt case marker was inherited from the proto-language, and a form without any case marker, referred to as the 'absolute', is still used in Old Georgian, beside the nominative in *-i*. The Common Kartvelian ergative marker is reconstructed as **-d/n*, primarily on the basis of evidence from Old Georgian and Svan (see the references above). This reconstruction was supported in Harris (1985: 79ff.) by the demonstration that the Old Georgian absolute is ordinarily used in the environments described independently in Greenberg (1978: 66ff.) for old noun forms dating from before the reanalysis and affixation of a determiner, while the nominative (with *-i*) is used in other contexts. The environments in which the absolute occurs include (a) in negation, (b) in predicate nominals, (c) in locative and temporal expressions, (d) in incorporated object constructions, (e) in compounds, and (f) with numerals.

At issue here is the loss of trapped morphology in the ergative; in (11) the material lost is underlined. As shown in (11), no morphology was trapped in the nominative, and loss of *gi* is due to other factors. In the ergative, the trapped suffix *-n* of the host noun was lost. If the input to grammaticalization was the proximal *a-man*, or remote *i-man*, then the vowel prefix was also lost; but here it is assumed for simplicity that the input was simply *man*.

Phonotactic constraints provide no explanation for loss of the noun's ergative case suffix in the development of the Georgian declension. On the assumption that no deictic element (*a-* or *i-*; see footnote 6) was present in the input to the change, the noun-determiner juncture would have been *n-m*, and this is not is ruled out on phonotactic grounds: *tan*m*iq'olebit* 'one after another', *vin*m*e* 'someone', *matgan*m*an* 'among them (ERG)'. (In each example, there is a morpheme boundary between the consonants in question, just as in the form at issue.) Thus, the case-determiner juncture is not a problem, but there is also the issue of the consonant clusters formed by the stem, the case, and the determiner. Modern Georgian clusters have been thoroughly studied (see Butskhrikidze 2002 and sources cited there). Consider the cluster that would have resulted in the ergative case of the noun *ʒeγl-* 'monument': **ʒeγl-n-man*. The cluster **γl-n-m* would probably not have been acceptable in Old Georgian, although *γl-m* poses no problem and occurs in the attested ergative form of this noun. It might thus be thought that problematic clusters such as **γl-n-m* would account for loss of the noun's case marker *-n-*. But consider, first, that there are also many vowel-final stems, and here no phonotactic problem arises; for example, in **ʒma-n-man* 'brother-ERG-ERG' we find the same relevant sequence as in *uban-man* 'district-ERG', where the former stem is V-final, and the latter *n*-final. If the trapped ergative case marker had been lost for phonotactic reasons, we would expect it to have been retained at least in vowel-final stems. Secondly, note that this cluster could have been resolved in several other ways, including the loss of any other consonant or the epenthesis of a vowel at any point. The simplest solution would have been not to use the

		SHORT FORM IN LPSL	SHORT FORM IN OCS	SHORT FORM +*j*-	LONG FORM IN LPSL	LONG FORM IN OCS
NOM SG	M	starŭ	starŭ	starŭ + jĭ	starjĭ	staryi, starŭi [starŭjĭ]
	N	staro	staro	staro + je	staroje	staro[j]e
	F	stara	stara	stara + ja	staraja	staraja
ACC SG	F	staro	staro	staro + jo	starojo	starojo
GEN SG	M/N	stara	stara	stara + jego	starajego	stara[j]ego, staraago, starago
INST PL	F	starami	starami	stary + jimi	staryjimi	stary[j]imi
LOC PL	F	staraxŭ	staraxŭ	stary + jixŭ	staryjixŭ	stary[j]ixŭ

Table 2
Short- and long-form adjectives in LPSl and OCS

basic *man*, shown in (11), but *a-man* (the proximal form), which also occurred as a definite article in Old Georgian. Thus, we conclude that in Georgian, phonotactic constraints cannot fully account for the loss of the trapped ergative suffix.

4. Slavic long-form adjectives

It is well known that Russian and other Slavic languages distinguish between so-called short-form and long-form adjectives. In Modern Russian, both the long and short forms can be used predicatively (with a semantic difference), but only the long form can be used attributively before a noun. (See Babby 1973, Bailyn 1994 for more detail on the uses and meanings of these forms.)

In Proto-Slavic, adjectives were definite or indefinite. The long form is the reflex of the definite adjective, formed from the adjective with the anaphoric pronoun *j*-. The short form is the reflex of the simple, inherited form, which was inflected as an *ŏ*-stem nominal for masculines and neuters, and as an *ā*-stem for feminines. It is the long form that is of interest from the point of view of trapped morphology. Table 2 is based closely on Schenker (1993) but represents a consensus view; for example, see also Andersen (1998). Late Proto-Slavonic (LPSl) forms in the table are reconstructed, Old Church Slavonic (OCS) forms attested. (⟨y⟩represents the high central vowel, ⟨j⟩ a glide.) In the singular forms cited here, the trapped case markers, each a single vowel (*-ŭ*, *-o*, *-a*, *y*), were all maintained. But in the plural forms cited, the disyllabic case exponents, *-ami* and *-axŭ*, 'were replaced by *-y*, extended analogically from the GEN PL' (Schenker 1993: 91). The fact that only

https://doi.org/10.1017/S0022226706003902 Published online by Cambridge University Press

disyllabic morphemes were lost indicates that phonology was involved. Although a phonological criterion determined which morphemes were lost, the loss itself was not the result of a phonological process, in the sense that it cannot be attributed to changes that occurred in the same phonological environment elsewhere in the language. Thus, the loss of disyllabic case endings in the formation of long-form adjectives in Proto-Slavic supports our claim that such loss can be morphological, not phonological.

In fact, if we look beyond the examples in table 2, we see that all disyllabic case markers in the trapped position were 'replaced by *-y*', including also the neuter instrumental singular *-omĭ*, neuter dative and instrumental dual *-oma*, neuter dative plural *-omŭ*, and neuter locative plural *-ěxŭ*. Why would singular and dual forms be analogically replaced by a plural? This phenomenon in Proto-Slavic makes much more sense if it is viewed as part of a common morphological process.

5. Georgian perfects and pluperfects

Forms such as those in (12) and (13) are attested in Old Georgian.[7] These are evidentials and (plu)perfects, and are translated here with (plu)perfect verb forms.

Old Georgian perfect (*evidential I*)

(12)	1SG damalul var	1PL damalul vart	'we are hidden, we have hidden'
	2SG damalul xar	2PL damalul xart	
	3SG damalul ars	3PL damalul arian	

Old Georgian pluperfect (*evidential II*)

(13)	1SG damalul viq'av	1PL damalul viq'venit	'we were hidden, we had hidden'
	2SG damalul iq'av	2PL damalul iq'venit	
	3SG damalul iq'o	3PL damalul iq'vnes	

These expressions consist of the past participle, *damalul*, and in (12), present-tense forms of 'be', and in (13), past-tense forms of this auxiliary.[8] The forms of the auxiliary here contain the Old Georgian subject–agreement markers shown in (14) and illustrated in (15) in another tense of the same verb.

[7] This description is limited to so-called absolute (single-argument) unaccusative intransitives and does not include transitives, unergatives, or relative intransitives (with two arguments), all of which have synthetic forms and are not relevant to the discussion here. See Harris (1985: 286–315) and Georgian sources cited there for description of the other forms.

[8] In certain tenses, plural direct objects and plural subjects of unaccusatives also condition the marker *-en*, which reduces to *-n* when followed by a suffix of the form -VC. This plural suffix is seen in the forms in (13), but not in those in (12).

(14)	1SG	v-		1PL	v-	-t	Old Georgian
	2SG	x/h/s/Ø-		2PL	x/h/s/Ø-	-t	
	3SG		-s/a/o	3PL		-es/-ian/-n/etc.	

(15)	1SG	da**v**imale		1PL	da**v**imalen**it**	'we hid'	Old Georgian
	2SG	daimale		2PL	daimalen**it**		
	3SG	daimal**a**		3PL	daimaln**es**		

When the Old Georgian expressions in (12) and (13) fused into single words, the first and second person prefixes *v-* and *x-* were trapped. Those in (12) were not lost, as shown in (16).

Modern Georgian perfect (*evidential I*)

(16)	1SG	da**v**malul**var**		1PL	da**v**malul**vart**	'we are hidden, we have hidden'
	2SG	damalul**xar**		2PL	damalul**xart**	
	3SG	damalul**a**		3PL	damalul**an**	

Modern Georgian pluperfect (*evidential II*)

(17)	1SG	da**v**malul**iq'av(i)**		1PL	da**v**malul**iq'avit**	'we were hidden, we had hidden'
	2SG	damalul**iq'av(i)**		2PL	damalul**iq'avit**	
	3SG	damalul**iq'o**		3PL	damalul**iq'vnen**	

A striking feature of (16) and (17), in contrast to Old Georgian (12) and (13), is the introduction of the agreement prefix *v-* by analogy to the predominant pattern, including forms such as those in (15). Its position in Modern Georgian is after a preverb (here *da-*) and before the stem, just as in other verb forms. It is clear that this *v-* cannot have moved, since *v* is still in position in *-var*, and since no other *v* was present in the corresponding form in Old Georgian (12).

The second person marker, *x-*, in the auxiliary, *xar*, was trapped but not lost because in the course of the history of the language, *x-* ceased to be a productive marker and was restricted to this auxiliary. The whole of the second person form of the auxiliary, *xar*, has been reanalyzed as a person marker, occurring in these forms and in the present tense of a handful of verbs, e.g. *t'iri-xar* 'you are crying'. The first person form of the auxiliary has been reanalyzed in a parallel fashion, even though *v-* continues to be productive and is now found earlier in the same verb forms (*v-t'iri-var* 'I am crying'). The plural formant *-t* is added after the person markers *-var* and *-xar*.

In the change from (13) to (17), the prefix of the first person (singular and plural) auxiliary was trapped between the participle and the auxiliary itself. As (17) shows, this trapped *v-* was lost. The second person prefix *x/h/s/Ø-* was also trapped, but in the change from Old to Modern Georgian this prefix was lost generally, except in the present tense of the verb 'be', as still retained in (16). Thus the loss of second person *x/h/s/Ø-* is not a change of the kind

discussed here; but the loss of *v-* in this position is.[9] (Other differences between (13) and (17) are the results of changes that applied more generally, including loss of the pluralizer *-en* (see Harris 1985: 209–228) and replacement of some third person plural subject markers by *-nen* (Harris 1985: 222–226, 397f.).) The endings of (17) remain identical to the past-tense form of 'be', except with regard to the placement of *v-*.

It is likely that the difference between the outcome in (16) and in (17) is due to the reanalysis of *-var* and *-xar* as alternative markers of the first and second person. They occur in this function with a handful of verbs in the present tense (see Harris 1985 for further arguments to this effect).[10]

This attested loss of *v-* in (17) cannot be ascribed to phonological change. It can be seen in (16) that the sequence *l-v-* at a morpheme boundary presents no problem, and the same sequence occurs as well elsewhere, e.g. *ulvašebi* 'mustaches'.

We conclude once again that this attested loss of *v-* is a morphological, rather than a phonological, change.

6. Greek futures

The development of the Greek futures is complicated and involved many changes; but it is both well attested and well studied, and this makes it more than worth including. Here we primarily follow the analysis of Joseph & Pappas (2002), drawing also on Joseph (1983, 2003), and we use their hypothetical examples, each of which is supported by attested examples. As the present paper deals only with morpheme loss, we examine only one small part of the story and begin in the middle.

Greek developed futures of two types, as in (18) and (19).

(18) Θelo: grafo:
'I will write.'

(19) Θelo: (i)na grafo:
'I will write.'

Θelo: 'I want' had been the first person singular present-tense form of the verb 'want', and it continues to function in that way. *Grafo*: is a first person singular subject form of 'write', and *(i)na* is a conjunction and later marker of the subjunctive. According to Joseph & Pappas (2002), a third type then

[9] As far as we are aware, the chronology of the general loss of *x/h/s/Ø-* relative to the specific loss of *v-* in this environment has not been studied in the detail that the loss of Scandinavian case markers has been (see section 2.1).

[10] Two examples of such paradigms are as follows:

1SG	v-t'iri-var	'I cry'	mi-v-di-var	'I go'
2SG	t'iri-xar		mi- di-xar	
3SG	t'iri-s		mi- di-s	

developed with the invariant third person form, *θelei*, as in (20), probably because of pressure to 'eliminate redundant person marking' from (18) and (19), and for other reasons.

(20) (a) Θelei grafo:
'I will write.'
(b) Θelei na grafo:
'I will write.'

Joseph & Pappas (2002: 253f.) argue that *θelei* in (20) is reduced to *θé*, probably via *θel* (cf. (21)); and they point out that the use of *θel* is attested and is found today in some dialects, such as Cretan, with *θela grafo*: or *θala grafo*: 'I will write'.

(21) Θel na grafo:

From (21), then, are derived both these dialectal forms and the modern standard in (22).

(22) Θa grafo:
'I will write.'

As it stands, this analysis is consistent with the claims made in this paper, though (20) is perhaps not exactly what one would expect. Although this is surely the right analysis of the Greek data in general, there are problems, which Joseph & Pappas themselves discuss (2002: 254, 263). One is that the form in (20a) is not attested until the 16th century, long after it should be, according to this scenario. Why, then, do Joseph & Pappas believe that (20) intervened between (18)/(19) and (21)? Their answer to this is the following:

> We differ from Meillet here in that we take the invariant impersonal 3SG form of *θelo*: as the most direct starting point for *θa*, since *θelei* is much more likely to yield truncated *θel* than first person singular *θelo*: is, given that the loss of unaccented high vowels is common in Greek; [footnote omitted] moreover, there are difficulties motivating the reduction in longer forms such as first person plural *θelome*(*n*) *na grafome*(*n*).

But we have seen from the examples above that the loss of trapped inflection is common cross-linguistically, not as a phonological change, but as a morphological change; and this may be expected to occur with longer inflection as well as with shorter inflection. Note that the inflection in (20) is *-o*:, and in the longer *θelome*(*n*), it is *-ome*(*n*); thus, loss of the trapped inflection (with the loss of *i-* in *ina* noted by Joseph & Pappas) in this instance would yield exactly (21), the form they propose as the immediate predecessor of the modern forms. Thus, if our suggestion is correct, it solves one of the problems that Joseph & Pappas pose for their analysis.

	STANDARD SPANISH	SELECTED AMERICAN DIALECTS
1SG	voy a dormir	yo v(w)adormir
2SG	vas a dormir	tu va:dormir
3SG	va a dormir	el va:dormir
1PL	vamos a dormir	nosotros vamos a dormir
2PL	van a dormir	Uds. va^{n} a dormir
3PL	van a dormir	ellos va^{n} a dormir

Table 3
Loss of trapped morphology in selected American dialects of Spanish (data from Fleischman 1982: 116)

7. American Spanish 'go'-future

A change in progress is the adoption of a synthetic future paradigm based on 'go' in some Central American dialects of Spanish.[11] While standard Spanish shows the forms on the left in table 3, these dialects show those on the right.

The forms in standard Spanish contain forms of the verb 'go' with suffixes indicating person and number of the subject; they also contain the particle *a* and the infinitive, in this case *dormir* 'sleep'. In the dialects at issue, the trapped person-number suffixes have been lost in the singular forms, though traces remain; in the plural forms, the suffixes have not (yet) been lost. Fleischman (1982) points out that the innovative forms require the use of a subject pronoun, though other forms of Spanish drop this. These Spanish dialects provide a second example of loss of agreement morphology trapped between a proclitic and a host.

8. Scandinavian reflexive verbs

In Old Norse the reflexive suffix *-sk* could be added to any verb form. This is a reduced form of the reflexive pronoun *sik*, which was in use during the same period (Ottosson 1992, 1999; Enger 2003; Faarlund 2005). Both the full form and the suffix are still used in Modern Scandinavian, but with different syntactic and semantic functions (Enger 2002). Unlike the full pronoun, the suffix was used to represent all oblique cases and both 2nd and 3rd persons. When the suffix was added to the present tense of a verb, the 2nd or 3rd person present ending *-r* was lost.

[11] The dialects include 'colloquial (educated) speech of Panama City, dialects of rural Mexican Spanish, and a lower-class dialect of Salvadorean Spanish' (Fleischman 1982: 104). Other material in this section is from Fleischman (1982: 115f.). We are grateful to an anonymous *JL* referee for bringing this example to our attention.

(23) *Old Norse*

þú kallar 'you call'	–	þú kallar þik	>	þú kallask 'you call yourself'
hann kallar 'he calls'	–	hann kallar sik	>	han kallask 'he calls himself'

This rule of loss of a trapped morpheme was not general, however. It never applies to vowels, nor does it apply to the consonants /m/ and /ð/ of the 1st and 2nd person plural, respectively.

Like other morphemes we have looked at, the present tense affix *-r* of the Scandinavian reflexive verbs is also not lost for phonotactic reasons; the suffix *-sk* can be added to stems in *-r*, e.g. *barsk* 'carried him-/her-/itself', and the cluster *-rsk* is found in other words, e.g. *fersk* 'fresh (fem.)', *rómnersk* 'Roman (fem.)'.

We conclude that the loss of trapped morphology in a given language may be both phonologically and morphologically selective.

9. Andi demonstratives

We now turn from these attested examples and well-accepted shallow reconstructions to newly proposed reconstructions. Andi is a language of the Andian subgroup of the Nakh-Dagestanian (North East Caucasian) family of languages, unrelated to Georgian and the Kartvelian family. It is a general feature of this family that most nouns have a so-called dual-base declension. That is, one stem occurs in the absolutive case, and a second stem is the basis of all other case forms. The second, or oblique, stem is derived from the first in some instances by ablaut, but usually by means of a suffix, usually called the oblique formant.

	NOUN CLASS I	'THIS' CLASS I	NOUN CLASS II	'THIS' CLASS II	NOUN PLURAL CLASS II	'THIS' PLURAL CLASS II
ABSL	hek'a 'man'	ho-**v**	yošk'a 'woman'	ho-**y**	moč'i-l 'children'	ho-**y**-il
ERG	hek'a-š:-d:i	ho-š:-d:i	yošk'a-l-d:i	ho-l-d:i	moč'i-l-**i**-di	ho-l-**i**-di
GEN	hek'a-š:-u-**b**	ho-š:-u-**b**	yošk'a-l-ɬi	ho-l-ɬi	moč'i-l-**i**-ɬi	ho-l-**i**-ɬi
DAT	hek'a-š:-u-y	ho-š:-u-y	yošk'a-l-ɭ:iy	ho-l-ɭ:iy	moč'i-l-**i**-y	ho-l-**i**-y

Table 4

Declension in two classes in Andi (Cercvaʒe 1965: 127, 129, 151, 204, 205 (Andi proper = Upper Andi))

While Andi has old, inherited dual-base declensions, the Andi forms in table 4 are recent innovations; the oblique formants here are derived from the demonstrative 'this', as shown by the forms in the table. While the

https://doi.org/10.1017/S0022226706003902 Published online by Cambridge University Press

ABSL	c'a	'fire'	k'otu	'horse'	ima	'father'
ERG	c'a-di		k'otu-di		im-u-di	
GEN	c'a-lí		k'otu-lí		im-u-**b**	
DAT	c'a-y		k'otu-y		im-u-y	

Table 5
Declension of some nouns in Andi (Cercvaʒe 1965: 125, 124)

absolutive forms are unaffected by the demonstrative, the oblique cases of the noun are formed historically by suffixing the appropriate form of the demonstrative 'this'. Although only four cases are listed here, Andi has many more cases, and all others are affected in a way similar to the ergative, genitive and dative. The oblique formant chosen reflects the difference between class I, formed with *-š:*, and other classes, formed with *-l*; the same formant occurs in the demonstrative and in the noun of the corresponding class. This *-l* comes from *-l:* synchronically.[12] The demonstratives and innovative nouns share the following peculiarities throughout the Avaro-Andian subgroup, which aids in this reconstruction:

- these oblique formants (of nouns)/roots (of demonstratives) are *-š:* and *-l:*
- different oblique formants/roots are correlated with different grammatical classes.

In addition, in the Andian subgroup (but not in Avar), the genitive of class I – nouns and demonstratives alike – is formed by suffixing a class marker (CM) that corresponds to the class of the possessed noun head. This may be *-b*, as in the example, for a class III or IV head noun, or *-v* for a class I head, *-y* for a class II head, or *-r* for a class V head.[13] The genitive of nouns and demonstratives of other classes is marked by the invariant suffix *-li*. Thus, in Andian languages all class I nouns have the following peculiarity:

- the genitive of a class I noun/demonstrative is formed with a CM.

These three peculiarities of the noun declension do not show up outside the Avaro-Andian subgroup and are not attributable to Proto-Dagestanian but reflect Proto-Avaro-Andian or Proto-Andian developments or diffusion within these subgroups.

Andi also possesses a simpler declension with a single base, illustrated in the first two examples of table 5. The third example in table 5 illustrates one

[12] The forms are affected by the following changes: d→d:/š:___ and ɬd→ld: and ɬ:→l/___ɬ (see Cercvaʒe 1965: 84). ⟨ɬ⟩ represents a voiceless lateral fricative, ⟨ɬ⟩ a strong voiceless lateral fricative, and ⟨ɬ′⟩ a strong voiceless lateral affricate.

[13] Five classes are characteristic of Andi proper, or Upper Andi. Lower Andi has three classes, while the Rik'vani subdialect has six classes (Cercvaʒe 1965: 93–102).

	NOUN CLASS I	RECONSTRUCTION	NOUN CLASS II	RECONSTRUCTION
ABSL	hek'a 'man'	*hek'a	yošk'a 'woman'	*yošk'a
ERG	hek'a-š:-d:i	*hek'-u-di-ho-š:-d:i	yošk'a-l-d:i	*yošk'a-di-ho-l-d:i
GEN	hek'a-š:-u-**b**	*hek'-u-**b**-ho-š:-u-**b**	yošk'a-l-łi	*yošk'a-lĺ-ho-l-łi
DAT	hek'a-š:-u-y	*hek'-u-y-ho-š:-u-y	yošk'a-l-ļ:iy	*yošk'a-y- ho-l-ļ:iy

Table 6
Reconstruction of case forms of *hek'a* and *yošk'a*

of several declensions with oblique formants different from those in table 4. Membership in declension classes is lexically determined.

Although the noun declensions illustrated in table 4 are an innovation, the nouns cannot have lacked declension before this innovation, since languages of this family generally have no undeclined nouns. Further, in the most closely related languages that lack the innovations described here, namely the Tsezian languages, the nouns cognate to *hek'a* 'man' are declined according to one of the patterns in table 5. For example, Tsez *žek'u* 'man, person' is cognate to Andi *hek'a* (Gudava 1979: 76) and is declined with a single base (Alekseev & Radžabov 2004: 120). It is likely that one of these patterns was used in the declension of *hek'a* 'man' in Pre-Andi.

We can tentatively reconstruct the forms in table 6, while recognizing that it is possible that at an earlier stage these nouns had no oblique formant or had a different oblique formant, such as *-i*, *-o*, *-ob*, which also occur in this function today. For 'woman', we have reconstructed forms following the declension of the first two nouns in table 5, which have no oblique formant, **yošk'a*, **yošk'a-di*, **yošk'a-lĺ*, **yošk'a-y*. This is the simplest declension pattern, and it is likely that **yošk'a* followed it because it is widespread for vowel-final nouns in the Tsezian languages. In the reconstruction in table 6, the noun is followed by the declined forms of the demonstrative 'this', shown in table 4. The source of the oblique formant may have been *ho-**v*** 'this, near the speaker' as reconstructed, but it could instead have been *he-**v*** 'that, near the addressee', *he-de-**v*** 'that, yon, far from speaker and hearer', *he-ge-**v*** 'yon, lower', *he-ļe-**v*** 'yon, higher', since these all share the same declension. We must also consider the possibility that what cliticized to the noun was a form that no longer exists, namely only the root of the demonstrative *– š*: or *l*: – with its suffixes, but lacking the deictic prefix. For 'man' we reconstruct forms following the declension of the third noun in table 5 above, *ima* 'father': *hek'a*, **hek'-u-d:i*, **hek'-u-**b***, **hek'-u-y* 'man'. Our confidence about the form of the oblique formant, **-u*, stems from the fact that 'man' is declined this way in some of the closely related Tsezian languages. The form of the genitive formant, a CM symbolized by ***-b*** (representing the full set of CMs), is also rather certain, since a CM is found

in the genitive for all members of class I in Andi. The demonstrative has effectively become an oblique formant.

The point we wish to make here is that the trapped morphology was lost. This included the noun case markers of the class II noun: **-di, *-lí*, **-y*, etc. For the class I noun, the trapped case markers **-di-*, ****-b-***, *-y-*, etc. were also lost. We also assume leveling of the stem form **hek'a* and loss of the oblique formant **-u*, but the latter loss may be considered part of leveling, not the type of loss considered here. For both nouns the deictic portion of the demonstrative (**-ho-*, **-he-*, **-he-de-*, **-he-ge-*, or **-he-ɬe-*), if present, was also lost. Because we cannot be certain of the presence of the deictic element in the input form, and because loss of the oblique formant could be considered part of leveling, we restrict our claim here concerning trapped morphology to the case markers – those listed above, as well as markers of additional oblique cases (see Harris 2005b for details).

The loss of case markers in Andi, like other losses described in this paper, cannot be explained as the result of phonological change. There are other words in which a CM occurs 'inside' another affix, such as the affective case form of 'man', *hek'a-š:-u-**b**-o*, or the CM agreement marker in *se-**v**-i k'otu* 'one horse'; thus we cannot assume automatic loss of the word-internal CM found in the genitive. Nor is there any phonotactic explanation for the loss of any of the trapped case markers. Thus, there are no phonological explanations available for the loss of case markers described here.

10. Zoque possessive constructions

Zoque, a member of the Mixe-Zoquean family spoken in Meso-America, has a set of case endings with local reference, as exemplified in (24a). The corresponding bare nouns are given in (24b).[14]

(24) (a) teʔ nüʔ-kukmü 'in the water'
teʔ mexa-küsi 'on the table'
kyopa-küsi 'on his head'

(b) nüʔ 'water'
mexa 'table'
kopa 'head'

The case endings *-kukmü* and *-küsi* and other similar case endings are bound forms. This is easy to ascertain for this language, since word boundaries are

[14] The Zoque data are from fieldwork by Jan Terje Faarlund. Examples are given in the common spelling now adopted by most of the community. The symbols have basically the same value as in (American) Spanish, thus ⟨j⟩ is /h/, ⟨x⟩ is /š/. The additional symbols ⟨ü⟩ and ⟨ʔ⟩ represent a mid central unrounded vowel and the glottal stop, respectively.

clearly marked phonologically. One criterion for this is metathesis: the semivowel /j/, spelled ⟨y⟩, always metathesizes with a following stop in the same word:

(25) tzakyüsi 'on the rope' <tzay küsi
kujkyüsi 'in the tree' <kujy küsi

There are also other phonological criteria, such as voicing (in some dialects) and stress placement, which yield the same result, but which we need not go into here.

In addition to these case endings, Zoque also has a set of simple case endings, including the locative *-i* and *-mü*. The longer, or secondary, case endings derive from nouns with the shorter endings. Thus, *kuk* 'middle' occurs as a noun with either of the two locative endings; *kuk-mü* has also been reanalyzed as 'in'. Similarly, *küs* 'body' (with the metaphorically extended meaning 'surface' or 'top') can occur in a locative case, and *küs-mü* has been reanalyzed as 'on'; see the example in (26).

(26) kimu küsmü
kim-u küs-mü
climb-COMPLETIVE top-LOC
'(S)he climbed to the top.'

The noun *küs* and similar nouns can still be used as relational nouns, (27), and as such they can also be used in possessive constructions, as in (28a, b).

(27) kuk-i kuk-mü 'in the middle of'
küs-i küs-mü 'on the body/surface of'

(28) (a) kumguʔyis kyuki
kumguy-ʔis y-kuk-i
village-GEN 3PS-middle-LOC
'in the middle of the village'

(b) teʔis kyüsmü ijtu teʔ kotzüjk
teʔ-ʔis y-küs-mü ijtu teʔ kotzüjk
DET-GEN 3PS-top-LOC be DET hill
'above it is the hill'

The contrast between 'in the middle of' and 'in' is expressed by the difference between the relational noun *kuk-mü* and the innovative case suffix derived from it. Similarly, the contrast between 'above, on the top of' and 'on' is expressed by the difference between the relational noun *küs-i* and the case suffix derived from it. The complex locative endings *-kukmü*, *-kuki*, *-küsmü*, *-küsi* exemplified in (24a) are the result of cliticization of the locative forms of the relational nouns. Indeed, it is only a short step from forms like the attested ones in (28) to the attested and cooccurring forms in (24). This step is shown in (29). In this process, both

https://doi.org/10.1017/S0022226706003902 Published online by Cambridge University Press

the genitive suffix *-ʔis* on the possessor and the person prefix *y-* on the relational noun are lost:

(29) (a) nüʔis kyukmü > nüʔ-kukmü 'in the middle of the water'
(b) mexaʔis kyüsi > mexa-küsi 'on the table'
(c) kyopaʔis kyüsi > kyopa-küsi 'on his head'

In (29a), it might appear that the full genitive, *-ʔis*, is not completely lost; but the remaining *ʔ* actually has a different source. It is part of the root, *nüʔ*, shown also in (24b).

An interesting feature about Zoque is that the independent relational noun and the bound form both co-exist in the language, as can be seen by contrasting (24a) with (28a, b).

In the related Mixean language Olutec, a similar situation obtains: Olutec does not have grammatical case endings (no genitive), but the possessed noun has a person prefix, as in Zoque. In Olutec the diachronic situation reconstructed for Zoque exists synchronically. Some relational nouns carry the person prefix, while others have been cliticized to the noun and have lost their prefix (Zavala 2000: 101f.). The least grammaticalized type is marked for both a possessor and a locative postposition. The relational noun in (30) has undergone semantic change but the syntax of the original oblique possessive construction is maintained.

(30) tükü ʔiniʔtzukpi
tük ʔi-niʔtzuk-pi
house 3PS-top-LOC
'on top of the house'

Other relational expressions have lost the possessor proclitic that intervened between the possessum (the relational noun) and the possessor. That is, the phrase is no longer a canonical genitive phrase.

(31) na:xküxmü
na:x-küx-mü
ground-surface-LOC
'on the ground'

In Zoque, as in the other languages we have examined here, the lost affixes would have been totally acceptable word-internally from a phonological point of view. The possessive prefix on the relational noun is underlyingly the semivowel /j/. This semivowel metathesizes with the initial consonant of the noun (in Zoque all words start with a consonant), and some consonants are palatalized. Thus on the surface the possessive prefix is not even a segment. Nonetheless it is lost in connection with the cliticization.

II. Discussion

In this paper we have shown, through attested and reconstructed examples from a variety of language families and types, that when two words are

juxtaposed in a way that could result in their fusing grammatically, trapping inflectional morphology between them, the trapped morphemes may be lost. We have not claimed that inflectional morphology in this position is always lost, and we know this to be untrue. In fact, we can observe that any one of at least three things may happen under these circumstances.

First, nothing may happen. Kabak (2006) documents in a number of language families (including, for example, Turkic and Kartvelian) the failure of postpositions to be reanalyzed as cases when they govern a case with an overt marker. Kabak argues that the presence of a case suffix on the noun 'protects' the adpositional phrase from grammaticalization.

Nevertheless, there are examples of postpositions being reanalyzed as cases even in the presence of a robust case suffix. Adams (1988) shows that in Proto-Tocharian, as well as separately in Tocharian A and Tocharian B, exactly this occurred. As one example, among those formed in Proto-Tocharian was the perlative, formed with *-*ā*, and this remained unchanged in Tocharian A. Klingenschmitt (1975) has shown that in Tocharian B the perlative plural, *-*ns*-*ā*, underwent boundary reanalysis, leaving the simple accusative plural ending in *-*n* and the perlative *-*sā*. Sapir (1936: 270, see also 263) identified Tocharian -*ā* as PIE **ad*. On the basis of this and other examples like it, we must conclude that the presence of a case suffix does not absolutely prevent grammaticalization of a postposition as a case. Nevertheless, it is true that when a noun and a postposition are juxtaposed, one possible outcome is that nothing happens.

A second possible outcome is that the trapped inflectional morphology remains. This has occurred, for example, in Skou, a language of New Guinea (Donohue 2003), in Kiranti languages of the Tibeto-Burman family (van Driem 1990, 1991, 1993, 1997), in North Ometo languages (Hayward 1998), and in Archi, Khinalug, Tsova-Tush and some other languages of the North East Caucasian family. In each of the languages named here, there are at least two sets of verbal agreement markers indicating the same NP argument in at least some circumstances. This can be illustrated with examples from Archi (Kibrik 1977: 128–130, 320, Corbett 1991: 108, 115f.), a language of the Lezgian subgroup of the Nakh-Dagestanian language family.

(32) **d**-as:á-**r**-ej-**r**-u-t:u-**r**
II-of.myself-II-SUFFIX-II-SUFFIX-SUFFIX-II
'my own'

(33) **w**-as:á-**w**-ej-**w**-u-t:u-**∅**
I-of.myself-I-SUFFIX-I-SUFFIX-SUFFIX-I
'my own'

The root in these examples is *as:á* 'of myself', while *t:u* is a suffix that forms adjectives. Each of the morphemes glossed with a Roman numeral is a marker of gender-class agreement representing the gender-class of the head, the possessed noun, which does not appear in the examples. As morphemes

Person/ number agree- ment prefixes	Tense- aspect- mood prefixes	Derivational prefixes	Verb root	Derivational suffixes	Tense- aspect- mood suffixes	Person/ number agreement suffixes

Table 7
Preferred order of morphemes

fused in Archi and the other languages named above, the trapped morphemes were NOT lost.

A third possible outcome is the loss documented here. We believe that loss of trapped morphology is the expected, default outcome when a clitic becomes an affix. We tentatively propose that in languages where two or more sets of agreement are tolerated, some special circumstance makes this possible, as it does in Skou (Donohue 2003) and Tsova-Tush (Harris 2005a).

We turn now to the question of why morpheme loss should indeed be the expected outcome, as claimed above. It has been recognized at least since Greenberg (1963) that there are universal preferences for morpheme order, and some of these can be summarized as in table 7. Synchronic aspects of this order have been much discussed, of course, with important contributions from Bybee (1985), Baker (1988), Rice (2000), and many others. From a diachronic point of view, many believe with Givón (1971) that morpheme order reflects order of grammaticalization. Yet it is implausible that inflection is always added 'last' in a language. We suggest that the loss of trapped morphemes is conditioned by the preference for maintaining these orders, as well as the preferred orders in substantives. Most of the examples we have described would have resulted in person-number agreement markers or case markers being trapped in a dispreferred position; in order to avoid this outcome, speakers have lost the trapped morphemes.

A different kind of change that has the same result is what Haspelmath (1993) calls 'externalization of inflection', the process by which inflectional affixes that are 'trapped' between other morphemes during the process of grammaticalization 'move' to an external position. Haspelmath's examples are limited to instances of inflected bases combined with an uninflected element, where loss of the trapped morpheme would result in complete loss of inflection in the affected item. Mithun (2000) extends the study of change in morpheme order by examining examples in which a derivational affix is reanalyzed as inflectional. While this does not necessarily change the actual relative order of morphemes, it can prevent a violation of the preferred morpheme order. Like the loss of trapped morphemes, these are important aspects of the way in which languages establish and maintain preferred morpheme orders.

Although the preferred solution for trapped morphemes is loss, we cannot truly predict when such morphemes will be lost, as in the examples examined in the sections above, and when they will be retained, as in Archi, examples (32) and (33) above, and Tsova-Tush. Consider, however, that assimilation is a specific process, relatively well understood; yet it is not possible to predict when it will apply. We cannot explain why assimilation does not apply everywhere where its structural description is met or why it begins to apply where it once had not. Similarly, we cannot fully explain why trapped morphemes are not always lost or why they are not lost earlier.

12. Conclusion

We suggest that relegating all loss during grammaticalization to erosion is vague and misleading. Recognizing a distinction between phonological erosion and the loss of trapped morphology is a first step in clarifying the changes that accompany the transition from independent word to clitic and on to affix. The loss of trapped morphology, we suggest, differs from erosion in at least the following ways:

- The morphological process typically targets complete morphemes, while erosion does not.
- Erosion is driven by prosody or phonotactic constraints.
- Erosion applies primarily in the word cliticizing (and its affixes), not in the host and its affixes.

In contrast, the loss of trapped morphemes targets complete morphemes. These can be a single segment, as were the Scandinavian genitive *-s* and the Georgian first person subject marker *v-*. But the trapped morpheme can be longer, such as Slavonic *-ami*, *-axŭ*, *-omĭ*, *-oma*, *-omŭ*, and *-ěxŭ* (see section 4 above), and the Greek person-number markers (see section 6 above). A typical example of erosion, on the other hand, is the process that reduced Old Norse *sik* REFLEXIVE to *-sk*, which targeted a single segment (section 8).

Erosion typically applies in unstressed syllables, as in Old Norse *sik*. Prosody does not drive morpheme loss, but this does not mean that the two do not interact. For example, it appears in the Slavonic examples that the case markers that were single vowels were not lost because they provided a suitable buffer or transition vowel, and that when the -VCV cases (listed in the preceding paragraph) were lost, they were replaced by *-y* because it too provided a transition vowel.

In the examples we are aware of, erosion is limited to the word cliticizing or its affixes. In contrast, the loss of whole morphemes applies to affixes of the host or the clitic, but is limited to those that are trapped.

In each section above we have argued that the loss of morphemes cannot be fully accounted for by the regular phonological processes of the language.

We have suggested that the loss of trapped morphemes is motivated by the need to establish optimal morpheme order. By losing inflection in a relatively internal position, the universally preferred morpheme order is reestablished.

Because this process targets whole morphemes, because it is not (entirely) part of regular sound change or a synchronic phonological process, and because it differs from phonological erosion, we conclude that it is an entirely morphological process.

REFERENCES

Adams, Douglas Q. (1988). *Tocharian historical phonology and morphology*. New Haven, CT: American Oriental Society.

Alekseev, M. E. & Radžabov, R. N. (2004). Tsez. In Job, Michael (ed.), *The indigenous languages of the Caucasus*, vol. 3: *North East Caucasian languages* (part 1). Ann Arbor, MI: Caravan Books. 115–163.

Andersen, Henning (1998). Slavic. In Ramat, Anna Giacalone & Ramat, Paolo (eds.), *The Indo-European languages*. London: Routledge. 415–453.

Babby, Leonard (1973). The deep structure of adjectives and participles in Russian. *Language* **49**. 349–360.

Bailyn, John F. (1994). The syntax and semantics of Russian long and short adjectives: an X′-theoretic account. In Toman, Jindřich (ed.), *Annual Workshop on Formal Approaches to Slavic Linguistics. The Ann Arbor Meeting: functional categories in Slavic syntax*. Ann Arbor, MI: Michigan Slavic Publications. 1–30.

Baker, Mark C. (1988). *Incorporation: a theory of grammatical function changing*. Chicago, IL: University of Chicago Press.

Booij, Geert (1994). Against split morphology. *Yearbook of Morphology 1993*. Dordrecht: Kluwer. 25–49.

Booij, Geert (1996). Inherent versus contextual inflection and the split morphology hypothesis. *Yearbook of Morphology 1995*. Dordrecht: Kluwer. 1–16.

Butskhrikidze, Marika (2002). *The consonant phonotactics of Georgian*. Ph.D. dissertation, University of Leiden.

Bybee, Joan (1985). *Morphology: a study of the relation between meaning and form*. Amsterdam: John Benjamins.

Cercvaʒe, Ilia (1965). *Andiuri ena* [The Andi language]. Tbilisi: Mecniereba.

Čikobava, Arnold (1939). Motxrobiti brunvis genezisisatvis kartvelur enebši [On the origin of the narrative case in the Kartvelian languages]. *Tbilisis Saxelmc'ipo Universit'et'is Šromebi* **10**. 2829–2838.

Corbett, Greville G. (1991). *Gender*. Cambridge: Cambridge University Press.

Donohue, Mark (2003). Agreement in the Skou language: a historical account. *Oceanic Linguistics* **42**. 478–498.

Driem, George van (1990). An exploration of Proto-Kiranti verbal morphology. *Acta Linguistica Hafniensia* **22**. 27–48.

Driem, George van (1991). Bahing and the Proto-Kiranti verb. *Bulletin of the School of Oriental and African Studies* **54.2**. 336–356.

Driem, George van (1993). The Proto-Tibeto-Burman verbal agreement system. *Bulletin of the School of Oriental and African Studies* **56.2**. 292–334.

Driem, George van (1997). A new analysis of the Limbu verb. In Bradley, David (ed.), *Tibeto-Burman languages of the Himalayas* (Papers in Southeast Asian Linguistics **14**). Canberra: Pacific Linguistics. 157–173.

Enger, H.-O. (2002). The story of Scandinavian -*s*(*t*) retold: grammaticalising a clitic to a derivational affix. *Folia Linguistica Historica* **23.1–2**. 79–105.

Enger, H.-O. (2003). Skandinavisk -*s*(*t*) en gang til: grammatikalisering fra klitikon til avledningsaffiks [Scandinavian -*st* once more: grammaticalization from clitic to derivational

affix]. In Faarlund, Jan Terje (ed.), *Språk i endring: Indre nosk språkhistorie*. Oslo: Novus. 29–56.
Faarlund, Jan Terje (2004). *The syntax of Old Norse*. Oxford: Oxford University Press.
Faarlund, Jan Terje (2005). From clitic to affix: on the history of Scandinavian reflexive verbs. In Nielsen, Hans F. (ed.), *Papers on Scandinavian and Germanic language and culture* (Special issue of *North-Western European Language Evolution* **46/47**). 53–72.
Fleischman, Suzanne (1982). *The future in thought and language: diachronic evidence from Romance*. Cambridge: Cambridge University Press.
Givón, Talmy (1971). Historical syntax and synchronic morphology: an archaeologist's field trip. *Chicago Linguistic Society* **7**. 394–415.
Greenberg, Joseph H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In Greenberg, Joseph H. (ed.), *Universals of language*. Cambridge, MA: MIT Press. 73–113.
Greenberg, Joseph H. (1978). How does a language acquire gender markers? In Greenberg, Joseph H., Ferguson, Charles A. & Moravcsik, Edith A. (eds.), *Universals of human language* (vol. 2). Stanford, CA: Stanford University Press. 47–82.
Gudava, Togo E. (1979). *Istoriko-sravnitel'nyj analiz konsonantizma didojskix jazykov*. Tbilisi: Mecniereba.
Harris, Alice C. (1985). *Diachronic syntax: the Kartvelian case* (Syntax and Semantics **18**). New York: Academic Press.
Harris, Alice C. (2004). Law and order: establishing and maintaining morpheme order through change. Presented at DIGS 8 (Diachronic Generative Syntax), Yale University, New Haven, CT.
Harris, Alice C. (2005a). Explaining exuberant agreement. In Eythorsson, Thorhallur (ed.), *Linguistic theory and grammatical change: the Rosendal papers*. Amsterdam: John Benjamins. [To appear.]
Harris, Alice C. (2005b). Reconstruction of declension in Andi. Ms., SUNY Stony Brook.
Haspelmath, Martin (1993). The diachronic externalization of inflection. *Linguistics* **31**. 279–309.
Hayward, Richard (1998). The origins of the North Ometo verb agreement systems. *Journal of African Languages and Linguistics* **19**. 93–111.
Imnaišvili, I. (1955). Nac'evari ɜvel kartuli [The article in Old Georgian]. *St'alinis Saxelobis Tbilisis Saxelmc'ipo Universit'et'is Šromebi* **61**. 249–276.
Joseph, Brian D. (1983). *The synchrony and diachrony of the Balkan infinitive: a study in areal, general, and historical linguistics*. Cambridge: Cambridge University Press.
Joseph, Brian D. (2003). Morphologization from syntax. In Joseph, Brian D. & Janda, Richard D. (eds.), *Handbook of historical linguistics*. Oxford: Blackwell. 472–492.
Joseph, Brian D. & Pappas, Panayiotis (2002). On some recent views concerning the development of the Greek future system. *Byzantine and Modern Greek Studies* **26**. 247–273.
Kabak, Bariş (2006). An obstacle to the morphologization of postpositions. *Studies in Language* **30.1**. 33–68.
Kibrik, A. E. (1977). *Opyt strukturnogo opisanija arčinskogo jazyka*, vol. 2: *Taksonomičeskaja grammatika*. Moscow: Izdatel'stvo Moskovskogo Universiteta.
Klingenschmitt, Gert (1975). Tocharisch und Indogermanisch. In Rix, Helmut (ed.), *Flexion und Wortbildung: Akten der V. Fachtagung der Indogermanischen Gesellschaft, Regensburg, 9.–14. Sept. 1973*. Wiesbaden: L. Reichert Verlag. 148–163.
Lundström, Gudrun (1939). *Studier i nyländsk syntax* [Studies in the syntax of Nyland]. Stockholm: Norstedt.
Mač'avariani, Givi (1960). Brunebis erti t'ip'is genezisisatvis svanurši [On the origin of one type of declension in Svan]. *Tbilisis Saxelmc'ipo Universit'et'is Šromebi* **93**. 93–104.
Mač'avariani, Givi (1985). Brunebis zogierti sak'itxi svanurši [Several issues concerning declension in Svan]. *Iberiul-k'avk'asiuri enatmecniereba* **23**. 231–241.
Mithun, Marianne 2000. The reordering of morphemes. In Gildea, Spike (ed.), *Reconstructing grammar: comparative linguistics and grammaticalization*. Amsterdam: John Benjamins. 231–255.
Norde, Muriel (1997). *The history of the genitive in Swedish: a case study in degrammaticalization*. Ph.D. dissertation, University of Amsterdam.
Ottosson, Kjartan G. (1992). *The Icelandic middle voice: the morphological and phonological development*. Lund: Department of Scandinavian Languages, Lund University.

Ottosson, Kjartan G. (1999). From cliticisation to inflection. Ms., University of Oslo.
Rice, Keren (2000). *Morpheme order and semantic scope: word formation in the Athapaskan verb*. Cambridge: Cambridge University Press.
Šaniʒe, Ak'ak'i (1957). Umlaut'i svanurši [Umlaut in Svan]. In Šaniʒe, Ak'ak'i (ed.), *Kartuli enis st'rukt'urisa da ist'oriis sak'itxebi*. Tbilisi: Universit'et'i. 323–376. [First published 1925.]
Sapir, Edward (1936). Tibetan influences on Tocharian, I. *Language* **12**. 259–271.
Schenker, Alexander M. (1993). Proto-Slavonic. In Comrie, Bernard & Corbett, Greville C. (eds.), *The Slavonic languages*. London: Routledge. 60–121. [Reprinted in paperback 2002.]
Teleman, Ulf, Christensen, Lisa, Hellberg, Staffan & Andersson, Erik (1999). *Svenska akademiens grammatik* (vol. 3). Stockholm: Svenska akademien.
Vangsnes, Øystein A. (1998). Double definiteness and possessive constructions in an Eastern Swedish dialect. In Haukioja, Timo (ed.), *Papers from the 16th Scandinavian Conference on Linguistics*. Turku: Department of Finnish and General Linguistics, University of Turku. 425–439.
Yip, Moira (1998). Identity avoidance in phonology and morphology. In Lapointe, Steven G., Brentari, Diane K. & Farrell, Patrick M. (eds.), *Morphology and its relation to phonology and syntax*. Stanford, CA: CSLI Publications. 216–246.
Zavala, Roberto (2000). *Inversion and other topics in the grammar of Olutec* (*Mixean*). Ph.D. dissertation, University of Oregon.

Authors' addresses: (*Harris*)
Department of Linguistics, SUNY Stony Brook, Stony Brook, NY 11794-4376, U.S.A.
E-mail: alice.harris@stonybrook.edu

(*Faarlund*)
Institutt for lingvistiske og nordiske studium, Postboks 1102, Blindern, NO-0317 Oslo, Norway.
E-mail: j.t.faarlund@iln.uio.no