

# The diachronic externalization of inflection<sup>1</sup>

MARTIN HASPELMATH

## *Abstract*

*Diachronic changes in the order of affixes have not often been described in the theoretical literature, but they do occur commonly in one situation: when inflectional affixes are trapped in an internal position as the result of the grammaticalization and affixation of an uninflected element (e.g. a particle on interrogative and demonstrative pronouns, or a reflexive marker on finite verbs), they are commonly externalized. A number of instances of this type of change from various languages are discussed, and some constraints obeyed by such changes are formulated. A common striking feature of such changes is intermediate hybrid forms that show the inflection both in internal and in external position. While the motivation for this kind of affix order change is straightforward, the existence of intermediate double-marked forms means that the mechanisms of simple repositioning (as in syntax) and proportional analogy are insufficient to understand the externalization of inflection. Instead, I argue that pleonastic affixation (as in feet-s) is the mechanism that accounts for hybrid forms.*

*Lessons to be drawn for the theory of morphology are the following: (i) morphology is not simply word-level syntax; (ii) only a theory that makes use of preference principles (as opposed to absolute constraints) is successful in explaining this type of change; (iii) morphological theory must reckon with conflicts among principles and with local optimization; and (iv) I tentatively propose a principle of conservatism in diachronic change to explain why pleonastic affixation may occur even if analogy is a possible alternative.*

## **1. Introduction**

While the diachronic reordering of elements in syntax has been the subject of a lot of linguistic research in recent decades, the reordering of morpho-

logical elements has received much less attention. One obvious reason for this is that diachronic changes in affix order have not been observed very often. Comrie (1980: 85) remarks that “diachronically, there are numerous attested examples of change in basic word order (e.g. SOV to SVO in the development from Latin to the Romance languages), whereas change in the order of morphemes within a word is...quite exceptional.”

However, diachronic change in the order of morphological elements is not entirely unheard of. This paper examines one class of such affix reorderings, which can be exemplified by the Georgian indefinite pronoun *rame* ‘anything’. This word inflects for case according to three different inflection patterns, as shown in (1a)–(1c) (data from Vogt 1971: 44–46).

(1) Georgian indefinite pronoun *ra-me* ‘anything’

	a. older pattern	b. intermediate pattern	c. newer pattern
nom.	ra-me		ra-me
dat.	ra-s-me	ra-s-me-s	ra-me-s
adv.	ra-d-me	ra-d-me-d	ra-me-d
gen.	r-is-me		ra-me-s
instr.	r-iti-me		ra-me-ti

The indefinite pronoun *rame* is derived from the interrogative *ra* ‘what’ by means of the indefiniteness marker<sup>2</sup> *-me*. In (1a), this marker is suffixed to the case-marking suffixes, whereas in (1c), the case markers follow the indefiniteness marker *-me*. In addition, there is a third intermediate pattern (1b) where the case inflections occur twice, both before and after *-me* (following Vogt’s terminology, I call these HYBRID forms). All three inflection types are attested in modern Georgian, but it is clear from the evidence of earlier stages of Georgian that (1a) represents the oldest pattern, while (1c) is the newest pattern. In the older pattern, the inflectional affixes are internal to the suffix *-me*, and in the newer pattern, they are external to it. The type of change that occurred between (1a) and (1c), the EXTERNALIZATION OF INFLECTION, is the topic of this paper.

In section 2, I present more data of this type from various languages that show that affix reordering is a much less exotic type of change than one might have expected. In section 3, I formulate some constraints on affix reordering of this type. Sections 4 and 5 are devoted to the identification of the motivation and the mechanism of the change. I argue that word-syntactic and analogical approaches are insufficient to explain the whole range of phenomena, and that the hybrid forms provide a key to the mechanism: pleonastic affixation, which is amply attested elsewhere. In section 6, I discuss various accounts of affix pleonasm, and in section 7, I conclude by reviewing the theoretical lessons learned from the external-

ization of inflection, especially the need for preference principles or “soft” constraints.

## 2. Some cases of externalized inflection

Inflectional affixes may be externalized if they end up in an internal position where they are followed by a postfix (i.e. a postinflectional affix), or more generally, in a position between the stem and an extrafix (i.e. an extrainflectional affix).<sup>3</sup> Such extrafixes generally derive from earlier clitic particles by grammaticalization. In the following sections, we look at cases of grammaticalization of extrafixes in certain types of pronouns (2.1, 2.3) and in finite verbs (2.2), and finally we consider an analogous change in certain types of compounds (2.4).

### 2.1. Externalization in pronouns

The grammaticalization of invariant clitic particles as extrafixes occurs particularly often in two types of pronouns: demonstrative pronouns and interrogative pronouns. Demonstrative pronouns are often reinforced by the addition of deictic particles, as in Swedish *den här bok* lit. ‘this here book’, Italian *questo* ‘this’ from *\*ecco iste* ‘behold that’, Polish *tamten* ‘that’ from *tam* ‘there’ plus *ten* ‘this’, etc. Extrafixes on interrogative pronouns are often indefiniteness markers, as in the Georgian example in (1) above, or in Russian *kto-to* ‘someone’, Latin *qui-dam*, *quae-dam*, *quod-dam* ‘a certain’, Finnish *kuka-an* ‘anyone’, etc. Interrogative pronouns may also be reinforced by “emphatic” particles that give the question a different pragmatic nuance, such as English *whoever*, Georgian *ra-ya* ‘whatever?’, etc.

A particularly well documented case of externalization of case inflection comes from Icelandic. In Old Icelandic, the free-choice indefinite pronoun *huerge* ‘any, whichever’ was formed from the interrogative pronoun *huerr* ‘which?’ by the addition of the postfixed particle *-ge* (cf. also *huat* ‘what?’, *huat-ke* ‘whatever’, etc.). Example (2) gives the paradigm of *huerr*, and (3) gives the paradigm of *huerge*. In (3), some slots contain more than one form, all of which are attested in the Old Icelandic literature. All the Old Icelandic data are from Noreen (1970 [1923]).

#### (2) Old Icelandic *huerr* ‘which?’

	masculine	feminine	neuter
sg. nom.	huer-r	huer- $\emptyset$	huer-t
gen.	huer-s	huer-rar	huer-s

	dat.	hueri-om	huer-re	hueri-om
	acc.	huer-n, hueri-an	hueri-a	huer-t
pl.	nom.	huer-er	hueri-ar	huer-∅
	gen.	huer-ra	(= m.)	(= m.)
	dat.	hueri-om	(= m.)	(= m.)
	acc.	hueri-a	hueri-ar	huer-∅

(3) Old Icelandic *huerge* 'whichever'

	masculine	feminine	neuter
sg. nom.	†huerge	†huerge	†huer(t)ke
gen.	°huerskes, huerges	hueregrar	°huerskes
dat.	huerionge, hueregom	hueregre, †huerrigi	huerionge
acc.	†huernge, °huern(e)gan	huerega	†huer(t)ke
pl. nom.	huereger	hueregar, †hueriage	†huerge
gen.	hueregra	(= m.)	(= m.)
dat.	†huerionge, hueregom	(= m.)	(= m.)
acc.	huerega	hueregar, †hueriage	†huerge

The paradigm in (3) shows side by side old forms (marked by a † sign) and new forms (unmarked), as well as hybrid forms (marked by a ° sign). (The old forms show a few mostly evident sound changes, such as dat. sg. m. \**hueriom-ge* > *huerionge*, which need not concern us.) In the old forms, the postfix *-ge* still follows the case suffix; in the new forms, it has been replaced by the suffix *-(e)g-* that comes between the stem and the case suffix; and the hybrid forms show the inflectional suffix both in its original internal position and in the new externalized position (*huer-ske-s*, *huer-n-(e)g-an*).

The next example is also from ancient Germanic, but from demonstrative pronouns. Ancient Germanic had a simple demonstrative m. *sa*, f. *sō*, n. *þat*, which later became the definite article in West Germanic (English *the*, German *der*). This could be combined with a deictic particle *se* or *si*, which was originally a separate word (cf. Gothic *sai*, Old High German *sē* 'behold') and later became grammaticalized and affixed to the demonstrative, yielding a complex demonstrative (later the West Germanic proximal demonstrative, English *this*, German *dieser*). The oldest forms of this complex demonstrative, attested in Runic inscriptions, show only the postfix *-si* (Brugmann 1904: 62).

## (4) Runic complex demonstrative forms

sg.	m.	f.	n.
nom.	sa-si	su-si	þat-si
dat.	þaim-si		
acc.	þan-si		

While these forms show no externalization yet, the process is already quite advanced in Old English. Compare the Old English paradigms of the definite article *sē* and of the proximal demonstrative *þēs* in (5)–(6). The old English data are from Campbell (1959: 290–292).

(5) Old English definite article *sē* ‘the’ (originally simple demonstrative)

sg.	m.		n.	f.	pl.
nom.	<i>sē</i>		<i>þæt</i>	<i>sēo</i>	<i>þā</i>
gen.		<i>þæs</i>		<i>þære</i>	<i>þāra, þæra</i>
dat.		<i>þām, þām</i>		<i>þære</i>	<i>þām, þām</i>
acc.	<i>þone</i>		<i>þæt</i>	<i>þā</i>	<i>þā</i>
instr.		<i>þon, þy</i>			

(6) Old English demonstrative *þēs* ‘this’ (originally complex demonstrative)

sg.	m.		n.	f.	pl.
nom.	<i>†þē-s</i>		<i>þis</i>	<i>†þēo-s, þios</i>	<i>†þā-s</i>
gen.		<i>þisses</i>		<i>þisse</i>	<i>þissa</i>
dat.		<i>þissum</i>		<i>þisse</i>	<i>þissum</i>
acc.	<i>þisne</i>		<i>þis</i>	<i>†þā-s</i>	<i>†þā-s</i>
instr.		<i>†þy-s</i>			

Only the acc. sg. f., the nom. acc. pl., and the instr. sg. m. n. forms of the proximal demonstrative represent the old forms with the postfix *-s*. The nom. sg. m. and f. forms have adopted the stem-initial consonant *þ*, and the other forms are formed from a stem *þis(s)-*. This stem is apparently based on the variant *þis* of the gen. sg. m. n. *þæs*.

In Old High German, the process of externalization is even more advanced than in Old English. Most forms are regular, based on a stem *des-*. Only the gen. sg. m. n. is sometimes attested in its original form *†des-se*. However, a few archaic hybrid forms are attested: gen. sg. m. n. *°des-se-s*, acc. sg. f. *°dhea-s-a*, nom. acc. pl. n. *°dei-s-u* (Braune and Mitzka 1963: 249).

A similar process of externalization of inflection must have occurred in the Latin word *ipse* ‘self’ (Brugmann 1904: 81). In Classical Latin, only external inflection is found: nom. sg. m. *ipse*, f. *ips-a*, n. *ips-um*, nom. pl. m. *ips-i*, f. *ips-ae*, n. *ips-a*, etc. However, in preclassical Latin, several forms are attested that show the inflection before an invariant postfix *-pse*. This was originally suffixed to the demonstrative *is* (f. *ea*, acc. sg. m. *eum*, etc.).

(7) Preclassical Latin

(nom. sg. m.	<i>*is-pse)</i>	nom. sg. f.	<i>†ea-pse</i>
dat. sg. m.	<i>†eo-pse</i>		
acc. sg. m.	<i>†eum-pse</i>	acc. sg. f.	<i>†eam-pse</i>

In addition, hybrid forms are sometimes found that show inflection both internally and externally.

## (8) Preclassical Latin

		nom. sg. f.	°ea-ps-a
acc. sg. m.	°eum-ps-um	acc. sg. f.	°eam-ps-am
		nom. pl. f.	°eae-ps-ae

The Classical Latin stem *ips-* is apparently based on a dissimilated nom. sg. m. *ipse* < \**is-pse*.

In the light of these cases, the inflection of the Ancient Greek demonstrative pronoun *hoũtos* begins to make sense. Compare the paradigm of the definite article *ho* (originally a demonstrative, cognate with Old English *sē* in [4]) in (9) and the demonstrative in (10).

(9) Ancient Greek definite article *ho*

	m.	n.	f.
sg. nom.	ho	tó	hē
gen.	toũ	tēs	tēs
dat.	tōi	tēi	tēi
acc.	tón	tó	tēn
pl. nom.	hoi	tá	hai
gen.	tōn	tōn	tōn
dat.	toĩs	taĩs	taĩs
acc.	toús	tá	tás

(10) Ancient Greek demonstrative pronoun *houtos*

	m.	n.	f.
sg. nom.	hoũtos	toũto	hautē
gen.		toútou	taútēs
dat.		toútōi	taútēi
acc.	toũton	toũto	taútēn
pl. nom.	hoútoi	taũta	hautai
gen.		toútōn	toútōn
dat.		toútois	taútais
acc.	toútous	taũta	taútās

The stem variants *hout-/haut-/tout-/taut-* are distributed in such a way that they are best interpreted as hybrid forms resulting from an earlier combination of *ho* plus a deictic particle *ute* (there is independent evidence for the existence of such a particle; see the discussion in Brugmann 1904: 103–106). The most straightforward forms are the nom. sg. m. *hoũtos* (< \**ho-ute*) and the neuter forms *toũto* (< \**to-ute*) and *taũta* (< \**ta-ute*), where the inflectional suffixes *-os*, *-o*, and *-a* have a simply been added

to the particle *ut(e)* (the final *-e* of this particle was dropped before the vowel-initial inflections). In all the other forms, only the vowel *o* or *a* of the original stem survives (in order to understand the sg. f. forms, it is important to know that the  $\bar{e}$  goes back to an earlier  $\bar{a}$ ). In the Boeotian dialect of Ancient Greek, this variation of the stem shape has been eliminated, and the stem is *hout-* for all genders and cases: nom. sg. m. *hoũtos*, gen. sg. m. *hoũtō*, acc. sg. m. *hoũton*, etc. (Brugmann 1904).

This kind of externalization is by no means restricted to Indo-European languages or languages of a similar morphological type (as, for instance, Georgian; cf. [1]). For example, the Basque demonstratives *hau* 'this (proximal)', *hori* 'that (medial)', *hura* 'that (distal)' can be reinforced by an "emphatic" suffix *-xe*. This is sometimes attached after the case suffix and sometimes precedes it (data from Saltarelli 1988: 214–216). Example (11) gives the forms of *hau*; the other demonstratives behave quite analogously.

(11) Basque proximal demonstratives *hau* and *hauxe* (some singular forms)

abs.	<i>hau</i>	<i>hau-xe</i>
erg.	<i>hone-k</i>	<i>hone-xe-k</i>
dat.	<i>hon-i</i>	<i>hone-xe-ri</i> , <sup>4</sup> † <i>hon-i-xe</i>
gen.	<i>hone-n</i>	<i>hone-xe-n</i>
loc.	<i>hone-ta-n</i>	<i>hone-xe-tan</i> /† <i>hone-tan-txe</i> <sup>5</sup>

Besides the forms with the stem *honex-*, one also finds forms based on the stem *hauxe-*, such as the locative *hau-xe-tan* (Martin Haase, personal communication).

Another example comes from Yakut, a Turkic language of Siberia. The Yakut indefiniteness markers *ere* and *eme* normally follow the inflected forms of the interrogative pronoun. But according to Ubrjatova (1982: 202), in the colloquial language the particle *eme* can take case inflections in addition to the interrogative pronoun, as shown in (12).

(12) Yakut *kim eme* 'anybody'

	standard	colloquial hybrid forms
nom.	<i>kim eme</i>	° <i>kim eme</i>
acc.	<i>kim-i eme</i>	° <i>kim-i eme-ni</i>
abl.	<i>kim-ten eme</i>	° <i>kim-ten eme-tten</i>
loc.	<i>tuox-ta eme</i>	° <i>tuox-ta eme-te</i>

Finally, there is a particularly complex example from Georgian involving two suffixes that become "internalized" successively. The indefinite pronoun *vi(n)-ya-c* 'somebody, anybody' consists of the interrogative pronoun *vin* 'who?' plus two particles, *-ya* and *-c*. Originally the particle complex *-ya-c* follows the case suffixes. At the next stage, the case suffixes

follow *-ya* but are still not fully externalized because *-c* still follows them. Only at the third stage have the case suffixes become fully external. All three stages are found side by side in modern Georgian (Vogt 1971: 45).

(13) Georgian *vi-ya-c* 'anybody, somebody'

	1st stage	2nd stage	3rd stage
nom.	vi-ya-c	vi-ya-c	vi-ya-c
dat.	vi-s-ya-c	vi-ya-sa-c	vi-ya-ca-s
erg.		vi-ya-ma-c	vi-ya-ca-m
gen.	vi-s(i)-ya-c		vi-ya-ca-s
instr.		vi-ya-ti-c	vi-ya-ca-ti
adv.		vi-ya-da-c	vi-ya-ca-d

2.2. *Externalization in verbs*

Internalization of inflection is particularly common in indefinite and interrogative pronouns, but is not restricted to them. Several cases from verbal inflection have been reported in the literature. Burrow and Bhattacharya (1970) mention a case from Pengo (Dravidian) where an original perfect postfix *-na* becomes a root suffix and the person/number inflection is externalized (see also Bybee 1985: 40). Three different types of the perfect coexist in Pengo. In the old type, *-na* still follows the person-number inflection; in the hybrid type, the person-number inflection occurs in the original position and after the perfect suffix *-na*; and in the new type, person-number inflection is completely externalized.

(14) Past and perfect in Pengo (Burrow and Bhattacharya 1970; Steever 1984)<sup>6</sup>

	past 'see'	perfect (old)	perfect (hybrid)	perfect (new)
sg. 1.	huɽtaŋ	ʃhuɽtaŋna	°huɽtaŋnaŋ	huɽtanəŋ
2.	huɽtay	ʃhuɽtayna	°huɽtaynaŋ	huɽtanəy
3m.	huɽtan	ʃhuɽtanna	°huɽtannaŋ	huɽtanəna
3f.	huɽtat	ʃhuɽtatna	°huɽtatnaŋ	huɽtanət
pl. 1ex	huɽtas	ʃhuɽtahna	°huɽtahnaŋ	huɽtanəs
1in.	huɽtap	ʃhuɽtapna	°huɽtapnaŋ	huɽtanəp
2.	huɽtader	ʃhuɽtaderna	—	huɽtanider
3m.	huɽtar	ʃhuɽtarna	°huɽtarnaŋ	huɽtanər
3f.	huɽtik	ʃhuɽtikna	°huɽtiknaŋ	huɽtanik
3n.	huɽtiŋ	ʃhuɽtiŋna	°huɽtiŋnaŋ	huɽtiniŋ

Another case of affix reordering in a verbal paradigm has been reported from a Lithuanian dialect (Stolz 1989). The data here are rather scanty,



but they fit the general pattern very well. Lithuanian has a postfix *-s(i)* that marks reflexive verbs and was grammaticalized from an earlier reflexive pronoun (see also Haspelmath 1990: 43). In standard Lithuanian, this is always the last suffix and follows the person/number suffixes, but in the dialects forms exist where another copy of the person/number suffix follows the reflexive suffix (*meldžiuosiu* in [15]) or where the order is directly reversed (*sukasim* in [15]).

(15) Lithuanian nonreflexive and reflexive verb inflection

	nonreflexive	reflexive	dialectal forms
	('work')	('work for oneself')	
sg. 1.	dirb-u	dirb-uo-s	°meldži-uo-si-u 'I am praying'
2.	dirb-i	dirb-ie-s	
3.	dirba	dirba-si	
pl. 1.	dirba-me	dirba-mė-s	°suka-si-m 'we are turning'
2.	dirba-te	dirba-tė-s	
3.	dirba	dirba-si	

Similarly, some varieties of Spanish have forms like *siénte(n)sen* 'sit down (imperative, plural)' (*siént-e-(n-)se-n* 'sit-SUBJ-(3PL-)REFL-3PL'), where standard Spanish has *siéntense* (*siént-e-n-se* 'sit-SUBJ-3PL-REFL') (Carmen Pensado, personal communication). Again, a reflexive suffix grammaticalized from a reflexive pronoun becomes a postfix, and the inflection is then externalized.

### 2.3. Externalization of prefixes

So far we have only seen examples where the reordering concerns an inflectional suffix that is externalized, thereby internalizing a postfix. But an inflectional prefix may also be externalized, thereby internalizing an antefix. Naturally such cases are much rarer due to the general dispreference against prefixes and the more specific dispreference against case prefixes (Kahr 1976). Nevertheless, I have one example to offer: in Russian, two indefinite pronouns have prefixal indefiniteness markers: *ni-kto* 'nobody' and *koe-kto* 'somebody (*kto* = 'who?'). These indefiniteness markers are antefixes because they precede case prefixes (usually called "prepositions" in Russian, but their word-internal position in these indefinites is a strong argument in favor of considering them case prefixes), as illustrated in (16).

(16) Russian case prefixes on indefinite pronouns<sup>7</sup>

	'nobody'	'somebody'	'somebody'
		(prescriptive)	(external inflection)
nominative	ni-kto	koe-kto	koe-kto
"adessive"	ni-u-kogo	koe-u-kogo	u-koe-kogo
"ablative"	ni-ot-kogo	koe-ot-kogo	ot-koe-kogo
"comitative"	ni-s-kem	koe-s-kem	s-koe-kem

In the prescriptive norm, the indefiniteness markers always precede the case prefixes, but nonprescriptive forms with the case prefix outside of the indefiniteness marker *koe-* are very common (Es'kova 1989: 65–66).

2.4. *Externalization in compounds*

In the preceding examples, the element that was externalized was always a kind of "particle" (indefiniteness marker, emphatic deictic marker, perfect marker, reflexive marker). However, a very similar phenomenon is also observed with certain types of compounds. Let me give three examples:

(i) In German adjective–participle compounds like *weit-gehend* 'far-reaching', *gut-verdienend* 'high-income', lit. 'well-earning', *viel-geliebt* 'much-loved', it is originally the first member that is inflected for comparative and superlative, as in *weit-er-gehend* 'more far-reaching', lit. 'further-reaching', *mei-st-geliebt* 'most-loved'. However, in the contemporary language this inflection sometimes appears in internal position, on the second compound member, as in *weit-gehend-st* 'most far-reaching', and hybrid forms with inflection both on the first and on the second compound member are also possible: *weit-est-gehend-st* 'most far-reaching', *be-st-verdienend-st* 'highest-income', *mei-st-geliebt-est* 'most-loved'. These hybrid forms in particular are frowned upon, but they are not uncommon in colloquial styles.

(ii) An English example is the compound *sister-in-law* (and other *-in-law* words), whose prescriptive plural is internal, on the first compound member: *sisters-in-law*. However, in casual usage the externalized plural *sister-in-laws* is quite normal, and even the doubly-inflected hybrid form *sisters-in-laws* is used by some speakers.<sup>8</sup>

(iii) In Classical Greek, past tense is marked by a prefix *e-* (the "augment"), as in *keĩtai* 'is lying', *é-keito* 'was lying'. In compound verbs consisting of local adverb + verb stem, the past tense inflection is sandwiched between the adverb and the stem, as in *pró-keĩtai* 'is lying in front', *pro-é-keito* 'was lying in front'. In later Greek, the augment is

sometimes placed in external position, preceding the adverb: *e-pró-keito*, and double-marked hybrid forms are also attested.

What all these compounds have in common is that they arose fairly recently by grammaticalization of a syntactic phrase to a compound; that is, they are JUXTAPOSITIONS. The initial internal inflection reflects the older syntactic pattern, but as the expressions come to be felt as single words, speakers externalize the inflection.<sup>9</sup> This is completely parallel to the cases of derivational affixes arising from agglutinated particles that we saw earlier.

### 3. Constraints on the externalization of inflection

The morphological changes that we have reviewed in the preceding section have quite a few things in common. Clearly, the diachronic process of externalization of inflection is subject to interesting constraints.

#### 3.1. Unidirectionality

First, it is clear that the change is unidirectional. Internal inflection may be externalized, but external inflection is never internalized. Consider the partial inflectional paradigm of the two Lezgian (Nakho-Daghestanian) demonstrative pronouns *am* and *at'am* given in (17) (data from Haspelmath 1993).

(17) Lezgian demonstratives *am* 'that one', *at'am* 'yonder'

abs.	a-m	at'a-m
erg.	a-da	at'a-da
gen.	a-da-n	at'a-da-n
dat.	a-da-z	at'a-da-z
adess.	a-da-w	at'a-da-w

One can imagine that the stem *at'a-* could be reanalyzed by speakers as consisting of the root *a-* plus the "emphatic" particle *-t'a*. After all, there are many languages where demonstratives are made up of a demonstrative root plus an "emphatic" particle. Then one could imagine a hypothetical change by which this particle is externalized, giving rise to such forms as *\*\*am-t'a*, *\*\*ada-n-t'a*, *\*\*ada-z-t'a*, *\*\*ada-w-t'a*, etc.<sup>10</sup> However, such a change has never been attested and must in all likelihood be excluded for principled reasons.

Since much of grammatical change is unidirectional, the unidirectionality of the externalization of inflection should not be too surprising.

However, unidirectionality is well understood only in the case of grammaticalization changes (see Lüdtke 1980 for an explanatory theory), and it is not clear that the changes we are dealing with here have anything in common with grammaticalization. To be sure, their preconditions arise by grammaticalization: a former free word is grammaticalized to a particle and clitic and is then suffixed to another word. But the subsequent developments are what interests us here, and they cannot be subsumed under grammaticalization. Thus, the unidirectionality of the externalization of inflection is a nontrivial constraint that needs to be explained.

3.2. *Restriction to inflection*

Second, only inflectional but not derivational morphology is externalized. This can be demonstrated in cases where analogous conditions obtain for both inflectional and derivational affixes. Consider the case of the Georgian indefiniteness marker *-me* (example [1] above). This is added not only to interrogative pronouns that inflect for case, but also to various other interrogative words that are formed by derivational affixes, as shown in (18).

- (18) Georgian adverbial interrogative and indefinite pronouns
- |         |               |           |                  |
|---------|---------------|-----------|------------------|
| sa-d    | 'where?'      | sad-me    | 'somewhere'      |
| sa-idan | 'from where?' | saidan-me | 'from somewhere' |
| ro-gor  | 'how?'        | rogor-me  | 'somehow'        |

Forms with externalized derivation can easily be imagined, but they are not created by the speakers (\*\**sa-me-d*, \*\**sa-me-idan*, \*\**ro-me-gor*).

3.3. *Hybrid forms*

Finally, only the inflectional affix is doubled in hybrid forms, while the particle is never doubled. Thus, we get the development in (19), where the oldest stage has the order inflection–particle, the intermediate stage has the order inflection–particle–inflection, and the final stage has the order particle–inflection.

- (19)
- |              | old form    | hybrid form      | new form   |            |
|--------------|-------------|------------------|------------|------------|
| a. Georgian  | ra-s-me     | ra-s-me-s        | ra-me-s    | (cf. [1])  |
| b. Icelandic | *huer-s-ge  | huer-s-ke-s      | huer-ge-s  | (cf. [3])  |
| c. Latin     | e-a-pse     | e-a-ps-a         | i-ps-a     | (cf. [8])  |
| d. Yakut     | kim-ten eme | kim-ten eme-tten | —          | (cf. [12]) |
| e. Pengo     | hurta-ŋ-na  | hurta-ŋ-na-ŋ     | hurta-na-ŋ | (cf. [14]) |

Again, it is easy to imagine hybrid forms where the particle is doubled, giving rise to a sequence particle–inflection–particle, such as Georgian **\*\*ra-me-s-me**, Icelandic **\*\*huer-ge-s-ke**, Latin **\*\*i-ps-a-pse**, Pengo **\*\*hurta-na-η-na**. Such forms would serve just as well as intermediate forms between the old stage and the new stage, but speakers never create them.

Explanations for these three constraints will be attempted below (see section 4.3 for an explanation of sections 3.1 and 3.2, and section 6.5 for an explanation of section 3.3).

#### 4. The motivation of externalization

It is immediately clear that the data presented in section 2 cannot be explained by considerations of formal simplicity. From the point of view of formal simplicity, there is nothing wrong with the pre-Old Icelandic paradigm **\*huerr-ge**, in which the suffix **-ge** was simply added to the inflected forms of **huerr**. The fact that speakers chose to change this simple paradigm and live with a rather chaotic mixed paradigm such as (3) for an “intermediate” period of many generations shows that they did not value formal simplicity as highly as is suggested by some morphological theories that focus on simplicity. Clearly, some substantive motivation must be involved to explain why speakers go to the trouble of changing the order of inflectional affixes and particles, even at the risk of substantial additional complexity.

##### 4.1. *The inflection-outside-derivation principle*

We do not have to look far for a substantive principle that motivates the externalization of inflection. Most linguists would agree that something like the well-known principle in (20) is at work here.

(20) *The inflection-outside-derivation principle:*

A morphologically complex word is preferred if its inflectional affixes are further away from the root than its derivational affixes.

This principle can account for the externalization of inflection in the cases cited in section 2 because the particle that becomes external as a result of the change has most of the relevant properties of derivational categories (even if it would not normally be treated under word formation in reference grammars). In particular, it creates new lexemes with a significantly different meaning. Thus, the Germanic and Ancient Greek complex demonstratives are the main demonstratives of the languages,

contrasting sharply with the simple demonstratives, which have become definite articles. Latin *ipse* means ‘-self’, contrasting with *is* ‘that; he’. In the indefinite pronouns, the difference in meaning between the base word (the interrogative word) and the derived word is even more striking. The Lithuanian reflexive affix also forms new lexemes from nonreflexive verbs. Only the Pengo perfect is not clearly a derivational category, but it is certainly much closer to the derivational pole of the derivation–inflection continuum (Bybee 1985: ch. 4) than the person/number affixes, which are prototypical inflectional forms.

Notice that (20) is formulated as a preference principle, that is, as a “soft” constraint that tolerates exceptions if other principles are in conflict with it and override it. This is necessary in order to explain why a change in the order of affixes takes place: if (20) were formulated as an absolute or “hard” constraint, forms like Latin *e-um-pse* would simply be excluded by the theory, the prediction being that such words become ill formed once the particle (*pse*) becomes agglutinated to the word. (See 7.2 for more discussion of preference principles.)

The preference principle formulated in (20) is of course not new. The cross-linguistic generalization concerning the order of derivation and inflection goes back at least to Greenberg (1963: Universal 28): “If both the derivation and inflection follow the root, or they both precede the root, the derivation is always between the root and the inflection.” The changes described in section 2 show that this is not merely a descriptive statement,<sup>11</sup> but that a principle like (20) is in fact used actively by speakers (cf. Bybee 1985: 40).

The principle in (20) can be further generalized and explained. According to Bybee (1985: 33), the basic principle is diagrammatic iconicity: elements whose meaning is more relevant to the stem are preferentially closer to the stem because this position iconically reflects the semantic relation. Derivational affixes are generally more relevant to a stem’s meaning than inflectional affixes, so (20) can be subsumed under Bybee’s relevance ordering principle. Dressler et al. (1987: 7) also formulate an explanation for the peripheral position of inflectional affixes:

The peripheral position of inflectional formatives facilitates their processing through the effects of psychological primacy and recency and better outward indexicality towards the other parts of the sentence. Derivational morphemes are also stored more than inflectional ones and thus should be stored together with the stem/root.

This explanation is formulated in psychological terms, but it is by no means incompatible with Bybee’s more abstract account. A deeper discussion of these matters is beyond the scope of this paper. My main goal in

this subsection is to point out that we need a theory that incorporates substantive preference principles like (20) in order to understand the phenomenon of externalization of inflection. The theories of Bybee and Dressler et al. provide such principles, whereas theories where the inflection-outside-derivation preference follows from grammar-internal formal restrictions cannot account for the changes in question.

#### 4.2. *Local optimization*

But if forms like Georgian *ra-s-me*, Latin *e-a-pse*, Spanish *siént-en-se* with derivation outside inflection are dispreferred and tend to be eliminated by language change, why do they arise in the first place? This is where the notion of LOCAL OPTIMIZATION comes in. The form *ra-s-me* is morphologically dispreferred, but it arose as a result of a completely independent change by which the particle *-me* was grammaticalized as an indefiniteness marker on the interrogative pronoun (see Haspelmath 1991 for this type of change). Grammaticalization changes, just like phonological changes, have their own rationale. They create their own local optimization, and they “do not care” whether they also create dispreferred structures elsewhere. Thus, the externalization of inflection can be regarded as a response by speakers to remedy the local dispreferred structures created by grammaticalization.

#### 4.3. *The explanation of the constraints in sections 3.1 and 3.2*

The motivation identified here immediately accounts for two of the constraints on this change that were noted above in section 3. Inflection is only externalized but never internalized (section 3.1) because internalization of inflection would create dispreferred structures without any local optimization in a different part of the grammar. And externalization is restricted to inflection (section 3.2) because only this leads to an improvement with respect to the inflection-outside-derivation principle. A hypothetical change of Georgian *sa-d-me* ‘anywhere’ to **\*\*sa-me-d** would not make the result more preferred because both *-d* and *-me* are derivational suffixes.

### 5. **The mechanism of externalization**

Having identified the motivation for the externalization of inflection, let us now ask exactly how the goal of complying with principle (20) is

achieved. It turns out that the morphology-as-syntax view (section 5.1) and analogical approaches (section 5.2) cannot answer this question. Only the examination of a related phenomenon, affix pleonasm (section 5.3) will bring us closer to understanding the mechanism of the change.

### 5.1. *Morphological reordering is different from syntactic reordering*

A big temptation in the study of morphology is to regard morphology as just the continuation of syntax with different means: a kind of word-internal syntax, where the smallest units that are concatenated are not words, but morphemes. However, the phenomenon described in section 2 above shows that affix order change proceeds in a very different way from word order change.

5.1.1. *Doubling is restricted to morphology.* If the internalization of inflection only involved the reordering of morphemes, we would expect it to proceed much like word order change: initially, the order is AB; then there is an intermediate stage at which orders AB and BA occur, with BA gradually increasing in frequency; and finally, BA is the regular order. The mechanism that we have seen in the reordering of morphological elements is quite different. Rather than an intermediate stage with variation between AB and BA, we generally find an intermediate stage with hybrid forms, where the inflectional affix is doubled: AB > ABA (hybrid) > BA. This intermediate hybrid stage has no analog in word order change and suggests that a very different mechanism is at work here.

It could be objected that an intermediate stage with doubling can be found in syntax as well (Theo Vennemann, personal communication), as in the change from preverbal to postverbal negation in some Germanic and Romance languages (Vennemann 1974). In English, for example, the original preverbal negation (*ne V*) changed to postverbal negation (*V not*) via an intermediate “hybrid” stage (*ne V not*). While there are no doubt certain parallels between this situation and the hybrid morphology discussed in this paper, there are also important differences:

(i) The double negation in *ne V not* is clearly the result of a grammaticalization change: the new negator *not* (<*naught*) initially serves as an additional emphatic word but is later reduced to an ordinary negative particle, making the old negator *ne* superfluous. In contrast, hybrid forms in morphology do not arise by grammaticalization.

(ii) The word order change is apparently an accidental byproduct of



the change from *ne* to *not*, while the morpheme order change discussed here is the goal of the change. (Vennemann [1974] argues that the word order change of the negator is part of an overall change from dependent-head to head-dependent order and therefore not at all accidental, but this is very doubtful: negators do not seem to behave like other verbal dependents cross-linguistically; see Dryer [1992: 97–98].)

It appears that double marking as a mechanism of word order change may occur in situations of language contact. For example, in Udmurt (Finno-Ugric; Vaxrušev et al. 1974: 90), conditional clauses are signaled by a clause-final conjunction *ke* 'if', which may cooccur with the clause-initial conjunction *esli* 'if', borrowed from Russian (see Thomason 1987 for more examples of this type). Again, in such cases the goal of the change is not to change the word order, but to assimilate one language to another, so it does not invalidate the claim that word order change works differently from morpheme order change.

5.1.2. *Treatment of allomorphy.* Another argument against a word-syntactic view of the externalization of inflection is the way allomorphs are treated. For instance, the Basque dative suffix is *-i* after consonants and has the allomorph *-ri* after vowels. The nonexternalized form *hon-i-xe* (cf. [11]) shows the suffix *-i* after the consonant-final demonstrative root. If the reordering were equivalent to a rearrangement of morphemes, we might expect the suffix *-i* to show up in its original shape, yielding *\*\*hon-xe-i*. However, the actual form is *honexe-ri*, with the allomorph *-ri* that would be expected if the form were created anew from the stem *honexe-*, but not if *honexeri* were the result of a simple rearrangement of the material in *hon-i-xe*. Note that the allomorphy is phonologically conditioned but does not follow from phonotactic restrictions; that is, *\*\*hon(e)xei* and *\*\*ra-me-is* (or *\*\*ra-m-is*) are possible phonological words in Basque and Georgian.

I have encountered one case of allomorphy that is not phonologically conditioned; again, the evidence is incompatible with simple rearrangement of material: the ergative case of Georgian *vin* 'who?' is identical to the nominative case, as in a few other pronouns. Likewise, the ergative case of nonexternalized *vin-me* 'someone, anyone' is *vin-me*. However, the externalized paradigm, which is based on the new stem *vinme-*, has not only the genitive *vinme-s*, the instrumental *vinme-ti*, etc., but also the ergative *vinme-m*, just like regular (nonpronominal) stems. Here there are no phonological reasons whatsoever that would disfavor an ergative case *\*\*vinme* in the externalized paradigm (i.e. *\*\*vin-me-∅*, the externalized variant of *vin-∅-me*).

5.2. *Proportional analogy is insufficient*

One obvious alternative to a word-syntactic approach is a morphological theory based on analogy (e.g. Becker 1990). It is easy to see how classical proportional analogy can account for the reordering and avoid one of the problems that a word-syntactic approach runs into. Consider again the Georgian data in (1). The only thing needed to create the externalized paradigm in (1c) is a reanalysis of *rame* (i.e. *ra- $\emptyset$ -me*) as *ra-me- $\emptyset$* , and analogy with other roots ending in *-e*, like *mope* 'king'. The proportion *mope : rame = (gen.) mope-s : X* yields the genitive *rame-s*, and so on for all the other case forms.

Such an approach makes the correct predictions about the allomorphy facts just mentioned in the previous subsection. If the genitive *rame-s* is created by analogy with the genitive *mope-s*, rather than by rearrangement of the elements of *r-is-me*, we would not expect a form *\*\*r(a)me-is* in the first place.

However, proportional analogy provides no better way of accounting for the hybrid forms than a word-syntactic approach. If the externalized paradigm nom. *rame*, dat. *rame-s*, inst. *rame-ti*, etc., is based only on a reanalyzed nominative form, we have no explanation for intermediate hybrid forms like dat. *ra-s-me-s*, adv. *ra-d-me-d*, etc.

Another serious problem with analogical approaches is that they work only if one form of the nonexternalized paradigm can be easily reanalyzed as a form of a regular paradigm, as in the case of *rame* (which is formally not different from *mope*). But consider the Pengo perfect (cf. [14]). The hypothesized nonexternalized forms are 1sg. *hurta $\eta$ na* 'I have seen', *hurta $\eta$ na* 'you have seen', *hurta $\eta$ na* 'he has seen', etc. None of these forms could possibly be reanalyzed as something else. The problem here is the absence of a zero form in the paradigm. If there were a 3sg. zero form *\*\*hurta- $\emptyset$*  'he saw', with a perfect *\*\*hurta- $\emptyset$ -na* 'he has seen', then this *\*\*hurta- $\emptyset$ -na* could have been reanalyzed as *\*\*hurta-na- $\emptyset$*  yielding by analogy the new paradigm *hurta-na- $\eta$*  'I have seen', *hurta-na-y* 'you have seen', etc. If proportional analogy were the only mechanism involved, we would expect externalization of inflection to occur only in those paradigms where one form (e.g. the nominative case, or the 3sg. form) is zero. But this is not the case: pronominal inflectional paradigms in older Indo-European languages commonly lack a zero form, but this does not block externalization in Old Icelandic ([2]–[3]), ancient West Germanic ([4]–[6]), or Latin ([7]–[8]).

Notice furthermore that proportional analogy would predict that the form of the stem should always be the same as in the zero form. Again, this prediction is in conflict with the data. In Basque, for example, there

is a zero form *hau-∅-xe*, which could conceivably be the basis for a reanalysis as **\*\*hauxe-∅**. But the other externalized forms are not based on this stem: the ergative is *honexek* (cf. the simple demonstrative *honek*), not **\*\*hauxek**; the dative is *honexeri* (cf. *honi*), not **\*\*hauxeri**, etc. (It is true that forms such as locative *hauxetan* also seem to exist, but the point is that proportional analogy cannot explain the other forms.)

I should stress that this criticism of proportional analogy does not mean that I want to deny that analogy plays a role at all. In quite a few cases of externalization, there is a zero form and no hybrid forms are attested, so proportional analogy is the most likely mechanism. Examples are Russian *koe-kto* ([16]) and Lithuanian reflexive verbs ([15]). However, I am more interested in other cases that do involve hybrid forms.

### 5.3. Affix pleonasm as a mechanism of externalization

We saw in subsections 5.1–5.2 that the existence of hybrid forms is a major problem for word-syntactic and analogical approaches. Recall also that the third constraint of section 3 (only inflection may be doubled in hybrid forms), which I have not accounted for yet, also involves hybrid forms. Hybrid forms are thus crucial for our understanding of the phenomenon of externalization of inflection, and I therefore turn to the discussion of a phenomenon that is closely related to hybrid forms.

**AFFIX PLEONASM** is the term introduced by Paul (1920: 162) for a type of morphological phenomenon that does not involve reordering but is otherwise quite parallel to the hybrid forms. Affix pleonasm consists in the semantically vacuous addition of a transparent affix to a word that is already characterized for the morphosyntactic property expressed by this affix. Some examples of pleonastic affixation in diachronic morphology are given in (21).

#### (21) Some cases of affix pleonasm

- a. The Latin infinitive *esse* 'to be' (root *es-*, infinitive suffix *-se*) was augmented by the more common infinitive suffix *-re* in Vulgar Latin (cf. Italian *essere*).
- b. The older English plural suffix *-er* was augmented by the more common plural suffix *-en* in the word *childer*, yielding forms like *children*.
- c. English plural forms lacking a final *-s* are augmented by the more common plural suffix *-s*, yielding forms like *feets*, *childrens*, etc., in some varieties of English (e.g. African-American Vernacular English, according to Hock 1986: 190).

- d. The older Latin 3pl. suffix of the perfect, *-ere*, was augmented by the more common 3pl. suffix *-unt* in later Latin, yielding the suffix *-erunt*.
- e. The standard (prescriptive) Spanish 2sg. preterite suffix *-ste* is augmented by the more common 2sg. suffix *-s* in “virtually all dialects” (Janda and Sandoval 1984) in nonprescriptive usage; thus prescriptive *lava-ste* ‘you washed’, nonprescriptive *lava-ste-s* (cf. *lava-s* ‘you wash’).
- f. The English comparative form *less* was augmented by the more common comparative suffix *-er*, yielding the form *lesser* (which coexists with the older form *less*, with a differentiated meaning).
- g. The German female suffix *-ess* in (archaic) *Prinz-ess* ‘princess’ was augmented by the productive suffix *-in*, yielding the modern form *Prinz-ess-in*.

This is just a small selection of such cases, which could be enlarged almost indefinitely (see Paul 1920: 162–63; Plank 1981: 76–82; Janda and Sandoval 1984; Horn 1988; Thomason 1987 for more examples). As several authors observe (Paul 1920: 163; Thomason 1987: 300), pleonastic affixation occurs mainly when the original form is irregular or unproductive, and the pleonastic affix is always a regular and productive one. Evidently, speakers use affix pleonasm to improve irregular forms on the parameter of morphosemantic transparency.

The parallel between affix pleonasm and hybrid formation is evident: in both cases, a form that is not optimal on one preference parameter is improved by augmenting it, rather than changing it analogically. In hybrid formation the problem is internal inflection; in affix pleonasm, the problem is insufficient morphosemantic transparency. The “therapy” is quite similar: just like Latin *laudav-ere* ‘they have praised’ becomes *laudav-erunt*, not *\*\*laudav-unt* (and *childer* becomes *children*, not *\*\*child-en*, or later *\*\*child-s*), Georgian *ra-s-me* initially becomes a hybrid form *ra-s-me-s* (which is then replaced by the completely externalized *rame-s*). Thus pleonastic affixation provides a mechanism for the rise of hybrid forms, and hybrid forms then serve as intermediate forms before the completely externalized forms are introduced.

## 6. Accounting for affix pleonasm

### 6.1. *Affix pleonasm creates new problems*

Affix pleonasm can thus be identified as the more general phenomenon of which hybrid formation is a special case. But affix pleonasm itself

needs to be explained. I have claimed that speakers add pleonastic affixes in order to improve certain problematic forms. But pleonastic affixation creates new problems for the speakers.

First, it is uneconomical to mark a form twice for one morphosyntactic property, and there is certainly a universal dispreference against lack of economy in morphology. Some morphologists have even expressed this as a universal "hard" constraint, often formulated as the morphological version of the *elsewhere condition* (e.g. Kiparsky 1982; Anderson 1986). It is stated as follows in Anderson (1986: 4):

Whenever one rule is more specific than another in the sense that the forms subject to the first constitute a proper subset of those subject to the second, the application of the more specific rule precludes the later application of the more general, less specific one.

If, as is typically done, stems that are lexically specified for a feature are taken to behave like a more specific rule, than one predicts that forms like *feet-s* should be impossible. And if the output of morphological rules is also taken as a lexical entry, then forms like *child-ren-s* and *es-se-re* should be impossible.

Second, affix pleonasm leads to additional allomorphy, and this violates the universal preference for uniform coding. For instance, the Vulgar Latin double-marked infinitive *es-se-re* must have been reinterpreted soon as *esse-re*, with a new stem allomorph *esse-*. Similarly, Spanish *lava-stes* is likely to be reinterpreted as *lava-stes*, with a new 2sg. allomorph *-stes*.

Thus, affix pleonasm appears to create as many problems as it solves. Given the concept of local optimization that I invoked earlier (section 4.2), this is not necessarily a problem: preference parameters are often in conflict with each other, so improvement on one parameter can lead to worsening on another parameter. But in quite a few of the cases of affix pleonasm and hybrid formation, analogy seems to be readily available as an alternative mechanism that avoids the problems of pleonasm. Why is Latin *laudav-ere* not replaced by *\*\*laudav-unt*, why is German *Prinz-ess* not replaced by *\*\*Prinz-in*? Before venturing my own hypothesis in section 6.4, I discuss two viewpoints on affix pleonasm in sections 6.2 and 6.3.

## 6.2. *Affix pleonasm as blending?*

Affix pleonasm is sometimes treated as an instance of BLENDING. Paul (1920: 162–63) introduces affix pleonasm under the general heading of

“contamination,” which in his terminology includes blending. Hock (1986: 189–190) uses the form *feets* as his first example of blending. According to him, “blending consists in the development of a morphological ‘compromise’ between two forms with identical or similar meaning which are perceived as being in competition with each other.” In the case of *feet-s*, the two forms that are blended are *feet* and the analogical *foot-s*, and *child-r-en* is a blend of the old form *child-er* and the newer (Middle English) form *child-en*. Similarly, one could claim that hybrid form like Georgian *ra-s-me-s* result from a blend of the old form *ra-s-me* and the new form *ra-me-s*.

While there may be some similarities between prototypical lexical blending like *brunch* (from *breakfast* and *lunch*) and affix pleonasm, the identification of affix pleonasm with blending must be rejected.<sup>12</sup> First of all, while blending is a “notoriously non-systematic or sporadic” change (Hock 1986: 189), affix pleonasm (and hybrid formation in the externalization of inflection) is not at all nonsystematic. It serves a clearly identifiable goal and is subject to specific restrictions (cf. sections 3.3, 5.3).

Second, an analysis of pleonastic affixation as blending presupposes the existence of analogically created forms with which the original forms can then be blended, like *foots* in Hock’s example. However, very often such analogical forms are not only unattested, but also impossible. For instance, the form *esse-re* can be explained as a result of blending only if an analogical form *\*\*es-re* existed at some point. However, this form cannot have existed because it is phonotactically ill formed. And as far as our hybrid forms are concerned, we saw in section 5.2 that proportional analogy is incapable of creating a regular form when there is no zero form that can be taken as the basis of the reanalysis. For example, the hybrid form Pengo *hurtaŋnaŋ* cannot be a blend of the old form *hurtaŋna* and the new form *hurtaŋaŋ*, because this new form cannot have arisen by analogy.

The hybrid forms cited in section 2 are clearly intermediate forms, not later forms than the regularized forms with external inflection. This is clear from the historical record in many cases, and also in modern cases where hybrid forms already exist, but completely externalized forms have not been created yet. For example, the Yakut hybrid forms in (12) are clearly very young forms, existing so far only in the colloquial language. The Yakut grammar mentions no analogical forms like *\*\*kim eme-ni* (‘anybody’, acc.), *\*\*kim eme-tten* (‘anybody’, abl.) that must be posited for an explanation of the hybrid forms *°kim-i eme-ni*, *°kim-ten eme-tten* in terms of blending.

6.3. *Phonologically motivated opacity?*

One reason why a morphological operation may be reapplied with no semantic effect (resulting in some kind of affix pleonasm) is that integrative phonological processes have severely reduced the syntagmatic recognizability of its first occurrence. Such cases are discussed in Plank (1985), where it is shown that such phonologically motivated opacity is one of the main factors allowing reapplication of morphology. For example, in some Bavarian dialects the plural suffix *-v* may be added to plural forms that are phonologically opaque, for example to the plural form *bubm* 'boys' (yielding *bubmv* 'boys'), which is not easily recognizable as a plural due to the fusion of the original plural suffix with the stem-final consonant ( $bubv \pm n \rightarrow bubvm \rightarrow bubm$ ).

Is syntagmatic opacity that is due to integrative phonology also a factor that can account for our hybrid forms of section 2? Naturally there are cases where the attachment of the postfix leads to some phonological changes of the inflected form. For example, the Old Icelandic dat. sg. m. *\*hueri-om-ge* becomes *huerionge*, the Pengo lexcl. pl. *\*hurtas-na* becomes *hurtahna*, and the Latin nom. sg. m. *\*is-pse* is dissimilated to *i-pse*. However, the syntagmatic opacity caused by these phonological changes is clearly insufficient to account for all the observed hybrid forms. For example, the preclassical Latin hybrid forms *eumpsum*, *eapsa*, *eampsam*, *eaepsae* do not show any opacity, nor do the Georgian forms or most of the other forms of the Pengo perfect paradigm. Thus, phonological opacity cannot be a necessary condition for the creation of hybrid forms.

This does not, however, mean that phonological opacity plays no role in the externalization of inflection. There is at least one respect in which integrative phonology is important: the reanalysis of the internalized particle as part of the stem. In several of the cases cited in section 2, the final result of externalization is that the internalized element is part of the stem. Thus, the Latin pronoun *ips-* is no longer synchronically segmentable as *i-ps-*; the Old English demonstrative stem *biss-* is no longer segmentable as *bis-s-*; the Ancient Greek demonstrative stem *tout-/taut-* is no longer segmentable as *to-ut-/ta-ut-*. In these cases, integrative phonology has struck, making the internal syntagmatic structure of these stems opaque.

6.4. *Conservatism as a motivating principle*

In section 6.1 above we saw that affix pleonasm creates its own problems. So why does it occur also in cases where analogy is a possible alternative?

In this section, I would like to suggest that there is an additional preference principle that is satisfied by pleonastically affixed and hybrid forms, and this is CONSERVATISM. Language change must be gradual, otherwise innovating speakers would not be understood by conservative speakers. So the innovations cannot be too radical. Other things being equal, speakers prefer a more conservative form that is closer to the earlier well-known form (e.g. the form used by noninnovating speakers of the same language). *Feets* is better than *foots* in that it is more similar to the older form *feet*, and *eumpsum* is better than *ipsum* in that it is more similar to the older form *eumpse*.

This is of course not a synchronic, cognitively based principle like the other preference principles that are commonly invoked in morphology (such as diagrammatic iconicity, morphosemantic transparency, economy, etc.). But since all languages are constantly changing, it is reasonable to assume that the mechanisms of language change play an important role in the way language is shaped.

In many cases (e.g. [20b]–[20d], [20f]–[20g]), the principle of conservatism must be invoked to explain why the analogical forms (\*\**child-en*, *foot-s*, \*\**laudav-unt*, \*\**littl-er*, \*\**Prinz-in*) are not used instead. In other cases, conservatism does not have to override another principle because alternatives are not available. Thus, Pengo speakers have no choice but to create hybrid forms like *hurta-η-na-η* if they want to get rid of inflection-internal forms like *hurta-η-na*. But the fact that *hurta-η-na-η* is so similar to the old, well-known form *hurta-η-na* (and is in this sense more conservative) probably helped “progressive” speakers to venture this innovation. The new form *hurta-na-η* cannot be created directly on the basis of the old form *hurta-η-na* because it is simply too different, too radical. Innovations can take only one step at a time, so hybrid forms like *hurta-η-na-η* or *e-um-ps-um* are necessary in order to get from *hurta-η-na* to *hurta-na-η*, or from *e-um-pse* to *i-ps-um*.

### 6.5. The explanation of the constraint in section 3.3

This scenario also explains the third constraint on the externalization of inflection formulated in section 3.3 above, the observation that in hybrid forms only the inflection is doubled, never the particle. Hybrid forms like *ra-s-me-s* or *hurta-η-na-η* arise by pleonastic addition of an inflectional affix to a form with internal inflection, resulting in a form that differs minimally from the earlier form. Hypothetical but nonexistent double-particle forms like \*\**ra-me-s-me*, \*\**hurta-na-η-na* are theoretically also possible intermediate forms between the old inflection-internal and the



new inflection-external forms, but there is no way to get these forms by simple pleonastic affixation. To get from *ra-s-me* to **\*\*ra-me-s-me**, or from *hurta-η-na* to **\*\*hurta-na-η-na**, a particle has to be inserted before the internal inflection. This would constitute as radical a deviation from the older forms as a direct reordering of the affixes and is therefore incompatible with the principle of conservatism. Only double-inflection hybrid forms like *ra-s-me-s* and *hurta-η-na-η* are possible because they preserve the syntagmatic structure of the older forms, just adding a new affix at the end.

This asymmetry nicely illustrates the fact that we are dealing here with externalization of inflection, not with internalization of a particle. Externalization is the primary event, and internalization is a secondary consequence of it. In hybrid forms, the inflection is already externalized, but the particle is not yet completely internalized. This primacy of externalization over internalization also justifies the title of this paper, which could not have been called "The internalization of particles."

#### 6.6. From hybrid forms to complete externalization

I have not said much so far about the change from hybrid forms to completely externalized forms. How does Georgian °*ra-s-me-s* become *ra-me-s*; how does Old Icelandic °*huer-s-ke-s* become *huer-ge-s*, how does Pengo °*hurta-t-na-t* become *hurta-na-t*? How do speakers get rid of the residual, nonfunctional internal inflection?

In some cases the answer is clear: by analogy. For example, the change from Latin acc. sg. m. *eumpsum* to *ipsum* can only be explained by analogy with the nom. sg. m. *ipse* (< \**is-pse*). Similarly, the regularized stem *hout-* in Boeotian Greek is based on the nom. sg. m. *houtos*. And the Old English dat. sg. m. *þissum* must be due to analogy with the gen. sg. *þisses* (= °*þis-se-s*). In other cases it is hard to tell because the data are inconclusive. For example, the precise origin of the new regular stem *huereg-* in Old Icelandic is not clear.

Thus, some details of the final "cleaning up" remain to be accounted for. However, this part of the change is clearly less surprising than the change that led to the hybrid forms, so a full account of it is perhaps less urgent.

### 7. Lessons for the theory of morphology

Let me now summarize what I see as the lessons for morphological theory that can be drawn from the phenomena studied in this paper and from the account which I propose here.

7.1. *Morphology is not simply word-level syntax*

First of all, the externalization of inflection once again shows that morphology cannot simply be treated as syntax at another level. This conception of morphology (represented, for example, by Selkirk 1982 and Lieber 1992) is not entirely misguided. After all, speakers continually create new morphology out of independent words by grammaticalization, and it is often unclear where syntax ends and where morphology begins. However, the phenomenon studied in this paper demonstrates the limitations of a word-syntactic approach. As we saw in section 5.1, the reordering of affixes shows no similarities with the reordering of words. Affix order change is one of the many areas where morphology is very different from syntax.

7.2. *Preference principles*

An important lesson to be drawn from the externalization of inflection is that morphological theory needs to incorporate preference principles, or “soft” constraints. We saw this in particular in the case of the inflection-outside-derivation principle (example [20]). This principle is evidently not an absolute or “hard” constraint that prohibits derivation outside of inflection, because forms with inflectional affixes inside the particle do occur, not just in the older stages of the languages cited in section 2, but also in many languages currently spoken. For example, in present-day Czech there is a demonstrative pronoun with internal inflection followed by a postfix *-to* (m. *ten-to* f. *ta-to*, n. *to-to*, etc.). Such structures are dispreferred according to the preference principle in (20), and we may expect Czech speakers to change them sometime. But they are not totally ruled out.

Morphological theories that centrally contain elements similar to preference principles have been proposed, among others, by Dressler et al. (1987), Bybee (1985), and Plank (1981) (and see Vennemann’s [1983] preference theories for phonology, and Jackendoff’s [1983] preference rule systems for lexical semantics and cognition in general).

By contrast, a theory where the inflection-outside-derivation principle is an absolute constraint cannot use this principle to explain why the change takes place. Such theories would incorrectly predict that once a structure with internal inflection arises, it will immediately become ill formed, that is, be lost from the language. Matters become even worse if the inflection-outside-derivation principle is not explained by grammar-external factors or stipulated, but if it follows from the organization of

the grammar, as in Anderson's (1992) theory. In such a theory, it would seem to be completely impossible to accommodate exceptions.

Another likely candidate for a preference principle is the *elsewhere condition* (cf. section 6.1). As we have seen, there are numerous cases of affix pleonasm that violate the elsewhere condition. Anderson (1986) discusses a few selected cases of affix pleonasm and tries to explain them away, suggesting a reanalysis of the earlier forms as noncomplex. Thus, a form like *feet-s* only makes sense in his theory if *feet* is somehow interpreted as a nonplural form.<sup>13</sup> This can perhaps be made plausible for this particular form (by appealing to considerations of local markedness; see Tiersma 1982), but there are enough other examples that cannot be disposed of in a similar fashion. For example, it is difficult to see how Latin *esse* 'to be' could have been interpreted as anything else but an infinitive form. And surely one would not want to say that hybrid forms like Georgian dat. *ra-s-me-s* are based on a dative *ra-s-me* that was reanalyzed as a nominative.

In view of these problems, it seems best to weaken the elsewhere condition in morphology to a preference principle. There is no doubt that there is a morphological dispreference against semantically vacuous affixation, and often the existence of certain nonderived stems blocks the creation of morphologically complex words with the same meaning (\*\**goed* because of *went*, \*\**stealer* because of *thief*). However, as the cases of affix pleonasm show, exceptions to it are possible under certain circumstances. Rather than simply abandoning the elsewhere condition, we should regard it as a preference principle that can be overridden if it is in conflict with other preference principles.<sup>14</sup>

Other morphological preferences that have briefly been mentioned above are morphosyntactic transparency (section 5.3) and uniformity (section 6.1).

### 7.3. *Conflicting preference parameters and local optimization*

Many preference parameters are in conflict with each other so that they cannot all be optimized at the same time. The motivation of language change is often the improvement of the system on one parameter (cf. Vennemann 1988; Wurzel 1985), but this may simultaneously lead to a worse situation on another parameter. However, this does not prevent speakers from carrying out such changes. They are blind to the further consequences of a change — all that counts is local optimization.

We have seen two examples of this: first, grammaticalization of certain particles on pronouns and finite verbs (as well as compounds of the

juxtaposition type) leads to dispreferred structures with internal inflection. Second, pleonastic affixation and hybrid formation improves the transparency of words, but at the same time leads to uneconomical and nonuniform coding. Again, this situation must be remedied, for example by analogical change.

#### 7.4. *Conservatism as a principle of language change*

Finally, I proposed that a principle of conservatism should be invoked to account for the existence of certain kinds of pleonastic affixation — certain types of change are motivated by speakers' attempts not to deviate too much from older patterns. This principle is admittedly rather ad hoc, but I see no other way of solving the problem of pleonastic affixation. It remains to be seen whether independent evidence for such a preference principle can be found.

*Received 5 January 1993*

*Freie Universität Berlin*

#### Notes

1. An earlier version of this paper was presented at the International Morphology Meeting, Krems (Austria), 7–9 July 1992. I am especially grateful to Jay Jasanoff, Carmen Pensado, Sally Thomason, Theo Vennemann, and Nigel Vincent for some useful suggestions. Correspondence address: Institut für Englische Philologie, Freie Universität Berlin, Gosslerstrasse 2-4, D-14195 Berlin, Germany:

##### *Abbreviations and notational conventions*

abs.	absolute case
acc.	accusative case
adess.	adessive case
adv.	adverbial case
dat.	dative case
erg.	ergative case
f.	feminine gender
gen.	genitive case
instr.	instrumental case
loc.	locative case
m.	masculine gender
n.	neuter gender
nom.	nominative case
pl.	plural number
REFL	reflexive marker
sg.	singular number
SUBJ	subjunctive
°	hybrid form, showing inflection twice
†	old form, showing original order of inflection and extrafix

\* reconstructed form

\*\* hypothetical but impossible form

2. See Haspelmath (1991) for the notion *indefiniteness marker* ('a grammatical element that turns an interrogative pronoun or a generic noun into an indefinite pronoun').
3. The term *postfix* is commonly used in Slavic linguistics, referring to postinflectional affixes like the reflexive marker *-sja/-s'* in Russian (cf. *ty moeš-sja* 'you wash [yourself]', *ona moeť-sja* 'she washes [herself]', etc.). The counterpart *antefix* (a preinflectional prefix) and the cover term *extrafix* were introduced in Haspelmath (1990: 29, 63).
4. The dative case suffix is regularly *-ri* after vowels.
5. The *ɪ* between *n* and *x* [ʃ] is due to a regular epenthetic process.
6. Steever's interpretation of the three types of the Pengo perfect differs from Burrow and Bhattacharya's. He proposes initial forms like *\*hurtaŋ (man)nay* (lit. 'he saw he is', a serial verb formation consisting of two finite verbs; cf. also Steever 1987: 78–83), which are optionally reduced by dropping the first or the second person-number suffix. If Steever is right, this would not be an instance of externalization of inflection, only of loss of double inflection. I am not competent to evaluate Steever's historical-comparative evidence, but a change from an original doubly inflected *\*hurtaŋ (man)nay* to an internally inflected *hurtaŋna* strikes me as very unlikely because it would violate the principle discussed below in (20).
7. Russian spelling always writes case prefixes ("prepositions") separately. When no case prefix intervenes, *ni-* is written together with the stem (*nikto*), and *koe-* is separated from the stem by a hyphen (*koe-cto*). Since this conventional spelling does not reflect linguistic reality, it is ignored in (16).
8. Thanks to Rand Valentine, Robert L. Davis, Paul Saka, Laurie Bauer, and the e-mail Linguist List for this information.
9. Stump (1991: 688) proposes a universal default rule (the "H-application default") to the effect that inflectional forms of endocentric words are formed from the head in the normal case. In Stump's account, forms like *sisters-in-law* and *pro-é-keito* follow from the universal default, whereas externalized forms like *sister-in-laws* and *e-prò-keito* require a special statement overriding the default. Stump does not specifically claim that his theory accounts for regularities of diachronic changes, but given his synchronic account, the observed diachronic tendencies are mysterious.
10. In this paper, I use the double asterisk to mark hypothetical, but ill-formed, nonexisting forms. The single asterisk is reserved for reconstructed forms that must have existed at some point but are unattested.
11. As a universal descriptive statement, it is of course too strong, because some cases of inflection inside derivation do occur. However, the fact that it has exceptions does not invalidate the generalization.
12. Paul (1920: 162) mentions a case that seems to be a true example of affix blending: in some varieties of German, forms like *rundlicht* 'roundish' are attested that contain the suffix *-licht*, a blend of the suffixes *-lich* and *-icht* (for other cases, cf. Plank 1981: 77–79).
13. See also Becker (1990: 24), who makes a desperate attempt to account for affix pleonasm by proportional analogy: "Es lässt sich zwar nur darüber spekulieren, was in den Köpfen von Kindern vorgeht, wenn sie solche Wörter bilden [i.e. *feets*], aber die Proportionalanalogie ist die einzige Erklärung, die plausibel erscheint. Von einem Kind, das die Form *feets* bildet, muss man wohl annehmen, dass es erstens die Form *feet* kennt und zweitens die Form *feet* nicht als den Plural 'Füsse' kennt, denn sonst würde es ja wohl *feet* sagen und nicht *feets*." Becker does not consider the possibility that the child knows that *feet* is plural but thinks it is not good enough as a plural and therefore adds the plural *-s*.

14. On the basis of examples like those in (21), Janda and Sandoval (1984) propose to abandon the elsewhere condition in morphology, claiming that "violations of the Elsewhere Condition among lexically-free morphological rules are, in fact, so widespread that they cannot legitimately be claimed to be 'marked' exceptions to some general Elsewhere Condition in morphology." If a view of grammar is adopted where preference principles are in conflict with each other, it is unnecessary to completely abandon a useful principle, even if it is often overridden by other prevailing preferences.

## References

- Anderson, Stephen R. (1986). Disjunctive ordering in inflectional morphology. *Natural Language and Linguistic Theory* 4, 1–31.
- (1992). *A-Morphous Morphology*. Cambridge Studies in Linguistics 62. Cambridge: Cambridge University Press.
- Becker, Thomas (1990). *Analogie und morphologische Theorie*. Studien zur Theoretischen Linguistik 11. Munich: Fink.
- Braune, Wilhelm; and Mitzka, Walther (1963). *Althochdeutsche Grammatik*, 11. Auflage. Tübingen: Niemeyer.
- Brugmann, Karl (1904). *Die Demonstrativpronomina der indogermanischen Sprachen: eine bedeutungsgeschichtliche Untersuchung*. Abhandlungen der philologisch-historischen Klasse der Königlich Sächsischen Gesellschaft der Wissenschaften 22: 6. Leipzig: Teubner.
- Burrow, Thomas; and Bhattacharya, S. (1970). *The Pengo Language: Grammar, Texts, and Vocabulary*. Oxford: Clarendon.
- Bybee, Joan L. (1985). *Morphology*. Typological Studies in Language 9. Amsterdam: Benjamins.
- Campbell, A. (1959). *Old English Grammar*. Oxford: Clarendon.
- Comrie, Bernard (1980). Morphology and word order reconstruction: problems and prospects. In *Historical Morphology*, Jacek Fisiak (ed.), 83–96. The Hague: Mouton.
- Dressler, Wolfgang; Mayerthaler, Willi; Panagl, Oswald; and Wurzel, Wolfgang U. (1987). *Leitmotifs in Natural Morphology*. Studies in Language Companion Series 10. Amsterdam: Benjamins.
- Dryer, Matthew (1992). The Greenbergian word order correlations. *Language* 68(1), 81–138.
- Es'kova, N. A. (1989). Formal'nye osobennosti nekotoryx predložnyx sočetańij s mes-toimennymi slovami. In *Russkie mestoimenija: semantika i grammatika*, T. A. Burlakova et al. (eds.), 60–68. Vladimir: Vladimirkij GPI.
- Greenberg, Joseph H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of Language*, Joseph H. Greenberg (ed.), 73–113. Cambridge, MA: MIT Press.
- Haspelmath, Martin (1990). The grammaticization of passive morphology. *Studies in Language* 14(1), 25–71.
- (1991). Zur Grammatikalisierung von Indefinitpronomina. In *Sprachwandel und seine Prinzipien*, Norbert Boretzky et al. (eds.), 103–125. Bochum: Universitätsverlag Brockmeyer.
- (1993). *A Grammar of Lezgian*. Berlin: Mouton de Gruyter.
- Hock, Hans Henrich (1986). *Principles of Historical Linguistics*. Trends in Linguistics. Studies and Monographs 34. Berlin: Mouton de Gruyter.
- Horn, Laurence R. (1988). Morphology, pragmatics, and the *un*-verb. *Eastern States Conference on Linguistics* 1988, 210–233.

- Jackendoff, Ray S. (1983). *Semantics and Cognition*. Current Studies in Linguistics Series 8. Cambridge, MA: MIT Press.
- Janda, Richard D.; and Sandoval, María (1984). *Elsewhere in Morphology*. Bloomington: Indiana University Linguistics Club.
- Kahr, J. C. (1976). The renewal of case morphology: sources and constraints. In *Working Papers on Language Universals* 20, 107–151. Stanford: Stanford University.
- Kiparsky, Paul (1982). Lexical morphology and phonology. In *Linguistics in the Morning Calm*, Linguistic Society of Korea (ed.), 3–91. Seoul: Hanshin.
- Lieber, Rochelle (1992). *Deconstructing Morphology*. Chicago: University of Chicago Press.
- Lüdtke, Helmut (1980). Auf dem Wege zu einer Theorie des Sprachwandels. In *Kommunikationstheoretische Grundlagen des Sprachwandels*, Helmut Lüdtke (ed.), 182–252. Berlin: de Gruyter.
- Noreen, Adolf (1970 [1923]). *Altnordische Grammatik*, 4. Auflage. Montgomery: University of Alabama Press.
- Paul, Hermann (1920). *Prinzipien der Sprachgeschichte*, 5. Auflage. Tübingen: Niemeyer.
- Plank, Frans (1981). *Morphologische (Ir-)Regularitäten*. Studies zur deutschen Grammatik 13. Tübingen: Narr.
- (1985). On the reapplication of morphological rules after phonological rules and other resolutions of functional conflicts between morphology and phonology. *Linguistics* 23, 45–82.
- Saltarelli, Mario (1988). *Basque*. London: Routledge.
- Selkirk, Elizabeth O. (1982). *The Syntax of Words*. Cambridge, MA: MIT Press.
- Steever, Sanford B. (1984). The evolution of the present perfect in Pengo. *Journal of the American Oriental Society* 104, 621–648.
- (1987). *The Serial Verb Formation in the Dravidian Languages*. Delhi: Motilal Banarssidass.
- Stolz, Thomas (1989). Zum Wandel der morphotaktischen Positionsregeln des baltischen Reflexivzeichens. *Folia Linguistica Historica* 9(1), 13–27.
- Stump, Gregory T. (1991). A paradigm-based theory of morphosemantic mismatches. *Language* 67(4), 675–725.
- Thomason, Sarah G. (1987). Double marking in morphological change. In *Eastern States Conference on Linguistics 1987*, 296–305. Columbus: Ohio State University.
- Tiersma, Peter M. (1982). Local and general markedness. *Language* 58, 832–849.
- Ubrjatova, Elizaveta I. (ed.) (1982). *Grammatika sovremennogo jakutskogo literaturnogo jazyka*, vol. 1. Moscow: Nauka.
- Vaxrušev, V. M.; et al. (1974). *Grammatika sovremennogo udmurtskogo jazyka: Sintaksis složnogo predloženiya*. Iževsk: Udmurtija.
- Vennemann, Theo (1974). Topics, subjects, and word order: from SXV to SVX via TVX. In *Historical Linguistics: Proceedings of the First International Congress of Historical Linguistics, Edinburgh, September 1973*, John Anderson and Charles Jones (eds.), vol. 2, 339–376. Amsterdam: North Holland.
- (1983). Causality in language change: theories of linguistic preferences as a basis for linguistic explanations. *Folia Linguistica Historica* 4(1), 4–26.
- (1988). Language change as language improvement. In *Modelli esplicativi della diacronia linguistica*, Vincenzo Orioles (ed.), 11–35. Pisa: Giardini.
- Vogt, Hans (1971). *Grammaire de la langue géorgienne*. Oslo.
- Wurzel, Wolfgang Ullrich (1985). Morphologische Natürlichkeit und morphologischer Wandel: Zur Vorhersagbarkeit von Sprachveränderungen. In *Papers from the 6th International Conference on Historical Linguistics*, Jacek Fisiak (ed.), 587–599. Amsterdam: Benjamins.

