

Journal of Universal Language 4
September 2003, 27-59

When Old Paradigms Die, New Paradigms are Born: On the Eternal Cycle of Morphological Change and its Importance for Language Typology

Gerd Jendraschek*

Université de Toulouse Le Mirail

Abstract

All languages change all the time, and the changes concern different linguistic domains. The first half of the present article will focus on the verbal morphology of Basque, a language where synthetic and periphrastic verb forms co-exist. Diachronically, the periphrastic forms have replaced most of the synthetic forms. The

* I would like to thank Michel Aurnague, Andrée Borillo, and Alan King, as well as two anonymous reviewers of the Journal of Universal Language for advice and comments. Thanks also to Fabio Montermini and Franck Floricic for information and advice on section 4. I am also very grateful to the Gottlieb Daimler- und Karl Benz-Stiftung, who provided me with a research grant, and to my home institution, the "Equipe de Recherche en Syntaxe et Sémantique" at the University of Toulouse-Le Mirail. Most thanks are due to my Basque informants/friends/consultants Larraitz Aranburu, Maria Jesús Lasa Barandiaran, Xabi Chabagno, Ramuntxo Christy, Maider Duguine, Ximun Duguine, Arantxa Eyherabide, Igor Herrarte, Ainoa Labeyrie and Asier Otermin, who are the true authors of much of this paper.

comparison of northern and southern varieties leads to interesting conclusions. Apart from obvious lexical differences, we can observe differences in the morphosyntax of the southern and northern standard, as speakers from the South use more morphological, synthetic forms, whereas northern speakers prefer analytical forms. Although the loss of synthetic forms has been observed for centuries, the dramatic loss of morphological complexity in the North seems to be a recent phenomenon. This morphological erosion is gradual. The study of synchronic and diachronic variation in Basque reveals that morphologically more complex forms and those representing marked categories are eliminated first, whereas shorter and more frequent forms are retained longest. In the second part of the paper, it will be shown that the universality of the phenomena described for Basque can be confirmed by data from other languages. In conclusion, the different data demonstrate how condensation of frequent forms and paradigmatic harmonization of isolated forms are responsible for the eternal cycle of morphological change.

Keywords: morphology, variation, language change, language typology

1. Basque: Regional Variation and Standardization

Basque is a language isolate spoken by approximately 550,000 people, 80% of whom live in the Basque Autonomous Community, composed of the territories of Bizkaia, Gipuzkoa and Araba, which are located on the Iberian peninsula, on the Atlantic coast to the south of the Pyrenees. The percentage of speakers is highest in Gipuzkoa and lowest in Araba. Approximately 10% of Basque speakers live in 'Iparralde' (literally 'northern side'), composed of the territories of Lapurdi, Low Navarre and Zuberoa (from west to east), on the French side of the border. The remaining speakers live in Navarre, where the North is Basque-speaking and the South al-

most exclusively Spanish-speaking.

Traditionally, one distinguishes the central dialects of Gipuzkoa and Lapurdi from peripheral dialects such as Bizkaian or Zuberoan, where interintelligibility is difficult to impossible. The standardization process in the 20th century is mostly based on the central dialects, but has led to slightly different results in the North, where Basque is in a diglossic situation with the ‘high language’ French, and in the South, where Spanish has had, and still has, a big influence on the Basque varieties. We can thus speak of a northern and a southern standard. In addition, there is not only diglossia between Basque and the dominating Romance languages, but between standard Basque and local varieties as well. The former is used in the media, education and all interregional contacts, whereas the latter are confined to use at home and in a small geographic area.

Apart from obvious lexical differences, we can observe differences in the morphosyntax of the southern and northern standard. When trying to use a more standard-like variety, speakers often just replace phonological and morphological forms of their native variety by those of the standard. Interestingly, speakers from the South use more morphological, synthetic forms, whereas northern speakers prefer analytical forms. Although the loss of synthetic forms has been observed for centuries, the dramatic loss of morphological complexity in the North seems to be a recent phenomenon, as more conservative grammars (see e.g., Lafitte 1944) report many forms which are no longer used today by young speakers.

2. Periphrastic and Synthetic Constructions in Basque

A first glance at a comprehensive grammar of Basque is likely to frighten the reader as he discovers lots of tables with paradigms

of complex looking verb forms. He will be relieved to find out that most of these forms are no more in use, or confined to literary use. This concerns many of the synthetic verb forms. In modern Basque, most verbs have only periphrastic forms. This is illustrated on the example of the verb *i-kus-i* ‘to see’ whose inflected forms consist of a participle marked for aspect and an auxiliary marked for tense, person and mood.

(1) Periphrastic Verb Forms (Based on Trask 1997:104)

recent past or perfect	<i>ikusi dut</i>	‘I have seen it’
present habitual	<i>ikusten dut</i>	‘I see it’
Future	<i>ikusiko dut</i>	‘I’ll see it’
remote past or pluperfect	<i>ikusi nuen</i>	‘I saw it’
past habitual	<i>ikusten nuen</i>	‘I used to see it’
etc.		

However, the same verb had synthetic forms some centuries ago. In the form *dakusat* ‘I see it’, inflectional morphemes are directly affixed to the stem *-kus-* (Altuna & Azkarate 2001:152). In the earliest texts from the 16th century, there occur about sixty verbs with at least some synthetic forms (Trask 1998:319). Today, only 10 to 20 verbs, depending on the speaker’s age and native dialect, conserve synthetic forms. Let’s have a look at the verb *j-oa-n* ‘to go’. The form in (2) is the synthetic form, (3) is the corresponding periphrastic construction where *da* is the auxiliary.

- (2) Aitor etxe-ra **doa**
 Aitor house-ALL go:PRS.3.SG
 ‘Aitor is going home.’

- (3) Aitor etxe-ra **joa-te-n** **da**
 Aitor house-ALL go-NR-LOC PRS.3.SG

‘Aitor goes home.’

The distinction is not only morphological, but also semantic: the synthetic form has nowadays a progressive reading, whereas *joaten da* refers to an habitual activity. In some cases, the periphrastic form adopts a progressive reading when the synthetic form is lost, in addition to the habitual use it already has. In other words, the aspectual distinction can be neutralized. There may be an intermediate stage, where the synthetic form is abandoned, but the aspectual distinction maintained by the use of different periphrastic constructions (Alan King, p.c.). Particularly speakers from the North tend to avoid synthetic forms, and in some cases conflate progressive and habitual aspect (Michel Aurnague, p.c.), but for most speakers, the distinction remains important. Most verbs do not have synthetic forms anyway, so that (4) is aspectually ambiguous (Zubiri & Zubiri 2000:413). But compare (4a), which is rather static than progressive, with (5), which is truly progressive and explicitly marked as such by the aspectual particle *ari*.

- (4) Musika entzu-te-n dut
 music listen-NR-LOC PRS. 3.SG<1.SG
 (a) ‘I hear music (now).’
 (b) ‘I listen to music (habitually).’

- (5) Musika entzu-te-n ari naiz
 music listen-NR-LOC PROG PRS.1.SG
 ‘I’m listening to music (now).’

3. The Reduction of Basque Paradigms

As mentioned in the preceding sections, morphological simplification is more advanced in the northern Basque Country, at least

as far as younger speakers are concerned. However, this does not mean that all forms are conserved in the South, nor that all forms have been dropped in the North. It is interesting to see which forms drop first, and which forms are maintained longest. This is difficult to do in diachrony, but the variation can be easily observed in synchrony, as forms can be evaluated by speakers and their frequency counted in corpora. The direction of morphological erosion is not random, but goes from marked to less marked categories.

Forms of possibility constitute a good example to illustrate the gradual obsolescence of morphologically complex forms. In ancient times, the modal suffix *-ke* could be added to a synthetically inflected verb, as shown in (6), one of the rare contemporary examples.

- (6) Azerbaijan-dik Turkia-ko kostalde-ra
 Azerbaijan-ABL Turkey-MR coast-ALL
doa-ke-en olio-bide-aeraiki nahi
 go:PRS.3.SG-POT-REL oil-way-DET build want
 luke, Errusia saihes-tu-z
 HYP.POT 3.SG<3.SG Russia avoid-PTCP-INSTR
 ‘It would like to build the pipeline that **might go** from
 Azerbaijan to the Turkish coast, bypassing Russia.’
 (*Euskaldunon Egunkaria* 14/05/02)

Another possibility of suffixing *-ke* was to add it to the indicative auxiliary, thereby conferring it an epistemic reading (examples from Zubiri & Zubiri 2000:420).

- (7) Xabierr-ek egin du
 Xabier-ERG do PRS.3.SG<3.SG
 ‘Xabier has done it.’

- (8) Xabierr-ek egin **du-ke**
 Xabier-ERG do PRS.3.SG<3.SG-POT
 ‘Xabier **might/must have** done it.’

The epistemic indicative is a counter-example to the overall tendency whereby ancient forms are better conserved in the South. These epistemic forms are still used in the North, whereas most speakers from the South do not know them or consider them as archaic (c.f., Zubiri & Zubiri 2000:440-441).

When combined with the subjunctive radical of the auxiliaries, the suffixe *-ke* corresponds to ‘may’ or ‘can’. However, there is a problem with transitive constructions whose direct object is a first or second person. These two categories are marked with respect to the unmarked third person. Thus, the form *zaitzake*, though regularly constructed, has become obsolete, and is avoided by many speakers.

- (9) Hurrengo Sanfermin-etan zezen bat-ek
 next Sanfermin-LOC.PL bull one-ERG
 harrapa zaitzake
 catch PRS.SUBJ.POT
 2.SG<3.SG

‘At the next San Fermin, a bull may get you.’

(www.armiarma.com/andima/korr/korr0624.htm)

The situation is similar with indirect object constructions, where three participants are involved. As agreement concerns all three participants—subject, direct, and indirect object—these auxiliaries are morphologically quite complex. However, the direct object can only be a third person in such a construction, as shown in (10).

- (10) Eman diezazuket
 give PRS.SUBJ.POT:3.SG<1.SG
 IO:2.SG

'I can give it to you.'

The first part of these forms, *dieza-* is invariable, but it can be followed by *-zki-* to indicate a plural direct object. Only the affixes surrounding *-ke* can vary to indicate first or second person. (11) shows the internal structure of these auxiliaries.

(11) Internal Structure of Three-participant Auxiliaries

di	-eza	-zu	-ke	-t
direct object	-radical	-indirect obj.	-modal suffix	-subject
3.SG	2.SG		1.SG	

Looking at frequencies of different forms gives interesting insights as to the evolution of the paradigm. (12) is based on a Basque corpus of 4.658.036 words (www.euskaracorporusa.net).

(12) Percentages of Personal Categories in three-participant Auxiliaries (total=142)

	subject	indirect object	direct object
1.SG	10%	5%	-
2.SG	8%	10%	-
3.SG	42%	48%	91%
1.PL	20%	8%	-
2.PL	6%	1%	-
3.PL	15%	27%	9%
	100%	100%	100%

It appears that the subject position of three-participant constructions allows of most variation as 43% of subjects are first or second person. Variation is somewhat reduced for indirect objects where these categories account for only a quarter, although, semantically,

the indirect object corresponds most likely to recipients or beneficiaries which are prototypically human. As said before, direct objects can only be third person, and strangely, only a small percentage concerns plural direct objects. Whereas in more frequent auxiliaries plural direct objects account for a quarter, here they only account for a tenth. The low frequency of plural direct objects in three-participant auxiliaries would be difficult to explain from a functional point of view and it is more plausible to say that forms are avoided when they become too ‘heavy’ as it is the case with a form like *di-eza-zki-zu-ke-t* which is composed of six morphemes. Avoiding the plural means avoiding one more affix, and this position of the paradigm is therefore reduced to the third person singular, i.e. the unmarked category. It is interesting to see that some isolated examples with non-third person direct objects are reported (see Trask 1997:220-221), but are not even mentioned by comprehensive grammars of Basque. The situation is thus different compared to the transitive auxiliaries illustrated by (9). Though avoided by many speakers, they are mentioned in grammars and used even by some younger speakers in the South.

The use—and thus the frequency—of certain auxiliaries depends on the accumulation of marked categories. The more an auxiliary expresses marked categories—past rather than present tense, first or second rather than third person, etc.—the more reduced will be its frequency and the more likely speakers will avoid to use it. Forms theoretically possible but inexistant in copora are often in past tense or involve three participants. It is not surprising then that the most frequent form is the one which accumulates only unmarked categories, i.e., present tense, a single third person singular participant and unmarked mood. This form is shown in (13).

- (13) Etorr-i da
 come-PTCP PRS:3.SG
 ‘He/she has come.’

In the before-mentioned four-and-a-half million word corpus, *da* appears 63.782 times, which contrasts a lot with the 82 occurrences of *duke* (8), the 3 occurrences of *zaitzake* (9) and the 11 occurrences of *diezazuket* (10).

The modal forms discussed so far can be considered obsolete or confined to stylistically marked use. Which are then the forms with the modal suffix *-ke* which are still widely used, i.e. which have escaped the morphological erosion? In transitive possibility constructions (those of the type exemplified by *zaitzake* in (9)), forms with third person objects account for 99% and are still widely used in the South (14), but have become obsolete in colloquial northern varieties, where periphrastic constructions as in (15) are preferred instead.

- (14) Lehiaketa-n edonor-k har **dezake**
 competition-LOC anybody-ERG take PRS.SUBJ
 parte
 part.POT 3.SG<3.SG
 ‘Anybody **can** take part in the competition.’
 (<http://www.herriak.info/nafarroa>)

- (15) Oraidanik eros-te-n **ahal** **dira**
 from.now.on buy-NR-LOC can PRS.3.PL
 sar-tze-a-k
 enter-NR-DET-PL
 ‘Entrance tickets **can** be bought from now on.’
 (www.herriak.info/lapurdi)

This difference between regional varieties of Basque concerns the possibility forms of intransitive auxiliaries in the same way. (16) provides statistical evidence for the differences concerning the frequency of such forms in regional corpora (taken from www.herriak.info).

small bowl.’

(MONASTERIO Xabier 1990, *Ipuin beldurgarriak*)

At an earlier stage, the alternative between the two forms corresponded to an aspectual difference. The synthetic form was progressive, the periphrastic form was habitual. This implies that complex synthetic forms like *daramakio* were much more frequent than they are today. In present-day Basque, the choice of synthetic (20) vs. periphrastic forms

(21) is more a matter of register, since, as mentioned before, many synthetic forms are restricted to literary use.

The differences in use for northern and southern varieties (and the corresponding regional standards) are summed up in (22). The reader should be warned that speaker judgments on particular forms may vary depending on age¹, education² and dialectal background, so the table cannot offer more than a general picture. As for the passive knowledge northern Basque speakers have of certain forms, two factors seem to be responsible. On the one hand, they may be remnants of older stages. On the other hand, they are the result of increased contact with southern varieties, e.g., through the media or school education. But even those of the northern speakers who make use of the more synthetic potential forms (those of the type of (14)), use the present tense forms much more than the corresponding past tense forms, whereas southern speakers use present and past tense forms equally: in a questionnaire survey, eleven Basque speakers were asked to evaluate certain forms. The maximum result of ‘10’ means that the form is used frequently, whereas ‘0’ means that the form is never used. The average results were as follows:

¹ Older speakers use more synthetic forms.

² Basque-medium education generally comprises some ‘training’ in Basque morphology.

- a) northern speakers not using synthetic potentiality forms: 2,1 for present tense forms (the forms to evaluate were *dezaket, ditzaket, zaitetzke*), 1,3 for past tense forms (*nezakeen, nitzakeen, zintezkeen*);
- b) northern speakers using synthetic potentiality forms: 7,0 for present tense, 2,8 for past tense;
- c) southern speakers: 9,6 for present tense, 9,3 for past tense.

(22) Use of Synthetic Verb Forms in Northern vs. Southern Basque Varieties

North (in particular Lapurdi and Low Navarre)	construction	example	South (in particu- lar Gipuz- koa)
passive knowl- edge or stylisti- cally marked use	intransitive present potential:	<i>(joan) zaitetzke</i> 'you can (go)'	actively used forms
	transitive present potential:	<i>dezaket, ditzaket</i> 'I can ... it'	
	tr. past. pot.:	<i>nezakeen, nit- zakeen</i> 'I could ... it.'	
	synthetic present	<i>darama</i> 'he/she wears/carries'	
	intr. past pot.	<i>joan zintezkeen</i> 'you could (go)'	
	synthetic past	<i>zeraman</i> 'he/she wore/carried'	
no longer used and often un- known	indirect obj. pres. potential	<i>(gerta) dakioko</i> 'it can (happen) to him/her'	passive knowledge or stylisti- cally marked use
	indirect obj. past pot.	<i>(gerta) zekioko</i> 'it could (happen) to him/her'	
	non-third person object pres. poten- tial	<i>zaitzaket</i> 'I can ... you'	
	non-third person object past pot.	<i>zintzakedan</i> 'I could ... you'	

	three participant present potential	(<i>eman</i>) <i>diezai-okezu, diezazki-okezu</i> 'you can (give) it/them to him/her'	
	synthetic present with indirect object agreement	<i>daramakiot</i> 'I carry it to him/her'	
	non-third person subject in indirect object construction	<i>zatzazkit</i> 'you ... to me'	no longer used and often unknown
	three participant past potential	(<i>eman</i>) <i>zeniezazkioken</i> 'you could (give) them to him/her'	
	synthetic past with indirect object agreement	<i>neramakion</i> 'I carried it to him/her'	
	potential suffixed to lexical verb	<i>dagoke</i> 'he/she/it can be'	

Let's recall the particular case of the epistemic form *duke* 'he/she might' (see example (8)), which is still widely used in the North, but nowadays usually unknown in the South.

The general impression produced by (22) is that the loss of verbal morphology concerns primarily three classes of forms:

- highly synthetic forms (those with certain suffixes directly attached to the lexical verb, such as *dagoke* or *daramakio*);
- forms requiring an accumulation of several agreement affixes (when three participants are involved, e.g., *diezaiokezu*, *zeniezazkioken*);
- forms with non-third person participants in syntactic positions where these are uncommon (*zaitzaket*, *zatzazkit*, etc.).

There is a tendency to avoid the morphologically most complex forms, beginning with the less frequent, marked forms, which can

more easily be dispensed with, and then passing on to the more frequent forms. The complex forms disappear first from colloquial speech, but may survive in more formal, literary contexts, so that diachronic change leads to the emergence of stylistic connotations.

In summary, two counteracting tendencies with respect to language change can be observed in modern Basque. On the one hand, the language evolves from a synthetic type to an analytic type. This evolution is taking place since first contact with Romance languages, but is accelerating in modern times and its effects can be best seen in the speech of young Basques from the North where most of the synthetic forms have been eliminated. On the other hand, synthetic forms have been conserved to a much larger extent in the southern varieties, as for example in the Gipuzkoan variety, which, of all the dialects, has the largest impact on the emerging southern standard. With southern media covering the North as well, and standardization affecting Basque-medium education in the North, an increasing number of morphological forms formerly confined to the South enter into the northern standard. Typological evolution goes towards less synthesis in all varieties, whereas standardization goes towards conservation of synthetic forms, and maybe, reintroduction of some of these forms in the North where they might be used as stylistically marked variants.

4. A Typological Evolution

The reader may wonder why there is so much variation in the morphology of a language. Indeed, the changes we can observe in Basque take place in many other languages in a similar fashion. Before trying to explain them, let us first see some more examples from different languages.

It is well known that there is a diachronic evolution concerning the interplay between analytic and synthetic morphology. It goes as

follows: Analytic expressions merge to synthetic forms, which are then replaced by new periphrastic constructions. We will now have a closer look at this evolution and the different steps of which it consists.

The transition from analytic to synthetic morphology is the result of a gradual loss of autonomy of the components of the analytic construction. It starts like this: Once fully autonomous words which happen to be frequently juxtaposed become attracted to each other and lose their positional and prosodic freedom, i.e., coalesce. Moreover, they become obligatory in certain contexts. In Spanish (23a-b) and Italian (c-d), subject personal pronouns are optional and can appear in different positions with respect to the verb, depending on pragmatic factors. In French however (e-f), the pronouns have become clitics, i.e., they have become part of the verbal complex where they must occupy certain positions, hence the impossibility of French (f) as the equivalent of Spanish (b) and Italian (d).

- (23) a. yo lo hago
 I it do:PRS.1.SG
 b. lo hago yo
 it do:PRS.1.SG I
 c. io lo faccio
 I it do:PRS.1.SG
 d. lo faccio io
 it do:PRS.1.SG I
 e. je le fais
 I it do:PRS.1.SG
 f. *le fais je
 it do:PRS.1.SG I
 ‘I do it.’

As coalescence proceeds, clitics become affixes. Affixes can merge, either with other affixes or with the base. This is what hap-

pens to some of the French pronominal clitics in informal spoken French. As we have seen, the first step is the restriction of positional freedom. With time, the less variable morpheme (generally the grammatical morpheme) agglutinates to the more variable morpheme (the lexical base). If both morphemes are frequent and constitute a recurrent combination, idiosyncratic inflectional forms may emerge. This is the case of forms like French *je suis* “I am”, where agglutination gave rise to an unsegmentable form where person-marker and base onset have merged.

(24) Phonetic Evolution of the French Personal Agreement Marker *je*

$ʒə\text{-}s\ddot{u}i > ʒ\text{-}s\ddot{u}i > ʃ\text{-}s\ddot{u}i > ʃ\ddot{u}i \text{ je suis "I am"}$
--

The merger of pronominal forms and verb forms is recurrent in the languages of the world. In German, personal subject pronouns can stand to the left or to the right of the verb. When they stand to the left, they display “strong” forms, whereas to the right of the verb, they often appear as “weak” forms, at least in informal speech. The verb *haben* “to have” has strong and weak forms as well, which correspond to formal and informal speech, respectively. Combinations of strong pronouns and weak verb forms, and vice versa, are possible, though. The important point here is that the strong pronominal forms have the status of autonomous units, whereas the weak forms are suffixes. The two paradigms have the same origine, but evolved in different ways.

(25) Strong and Weak Personal Pronouns in German

Grammatical category	written form	strong pronom. forms	strong verb forms	weak verb forms	weak pronom. forms
PRS.1.SG	<i>ich habe</i>	ɪç	'hɑ:bə	hap	ɪç
2.SG	<i>du hast</i>	du	hast	hast	ə
3.SG.M	<i>er hat</i>	eɪ̯	hat	hat	e
3.SG.F	<i>sie hat</i>	zi			zə
3.SG.N	<i>es hat</i>	es			s
1.PL	<i>wir haben</i>	viɪ̯	'hɑ:bən	ham	və
2.PL	<i>ihr habt</i>	iɪ̯	hɑ:pt	hapt	e
3.PL	<i>sie haben</i>	zi	'hɑ:bən	ham	zə

You may have noticed that similar processes can be observed in English, too, as with *I'm*, *you'll*, *don't*, *ain't*, *won't* etc. It should be remarked that in English, it seems to be the auxiliary which is structurally reduced or compressed, whereas the pronouns seem relatively stable. What is common to both cases is the observation that recurrent forms are prone to structural reduction and coalescence.

The tendency for free pronouns to become clitics and eventually agreement affixes can be further illustrated by data from Basque (see Trask 1997:218): *Gu* “we” and *zu* “you” are free pronouns, but the same forms are also verbal agreement suffixes, as shown in

(26) and (27), where the pronouns are optional and used only for emphasis.

(26) [**Gu zu-ri**] eman dizki-**zu-gu**
 we you-DAT give:PTCP ABS.3.PL-IO.2.PL-ERG.1.PL
 ‘We have given them to you.’

(27) [**Zu gu-ri**] eman dizki-**gu-zu**
 you we-DAT give:PTCP ABS.3.PL-IO.1.PL-ERG.2.PL

‘You have given them to us.’

All this is about structural reduction and/or coalescence, the ultimate stage of which is fusion. On the cognitive level, becoming a ‘single processing unit’ is a prerequisite to fusion. As the label says, single processing units are not decomposable, i.e. they are not resolvable into their constituents. This refers to the cognitive treatment of these forms, not etymological analysis. Consider examples (28)-(31) from Turkish: Standard Turkish verb forms in progressive (28) and future (29) aspect are decomposed into (at least) three morphemes: root, aspect marker, and personal agreement marker. The agreement marker is zero for third person singular referents. A suffix for negation can be inserted between root and aspect marker.

(28) O gel-iyor, sen gel-m-iyor-sun
 D.3(SG) come-PROG(3.SG) 2.SG come-NEG-PROG-2.SG
 ‘He is coming, you are not coming.’

(29) O gel-ecek, sen gel-me-yecek-sin
 D.3(SG) come-FUT(3.SG) 2.SG come-NEG-FUT-2.SG
 ‘He will come, you will not come.’

As combinations of aspect markers and agreement suffixes are very frequent, they come to be cognitively treated as being a single unit. The consequence on the structural level is fusion which can be seen in colloquial Turkish, where (28) and (29) become processed as shown in (30) and (31), respectively (Götz 1991/1995).

(30) O gel-iyo, sen gel-m-iyon
 D.3(SG) come-PROG.3.SG 2.SG come-NEG-PROG.2.SG
 ‘He’s coming, you ain’t coming.’

- (31) O gel-cek, sen gel-mi-cen
 D.3(SG) come-FUT.3.SG 2.SG come-NEG-FUT.2.SG
 ‘He’ll come, you won’t come.’

Fusion of morphemes means that two signifieds become associated with a single signifier. These phenomena are known as cumulative or portmanteau morphemes. Segmentation into the two original morphemes becomes impossible, at least outside etymological analysis. In standard Turkish, the interrogative particle *mi* splits the verbal ending (32), but in the informal variety, speakers cannot insert the interrogative particle between the originally two morphemes, as they have become one (33).

- (32) Gel-ecek mi, gel-me-yecek mi-sin?
 come-FUT(3.SG) INT come-NEG-FUT INT-2.SG
 ‘Will he come, won’t you come?’

- (33) Gel-cek mi, gel-mi-cen mi?
 come-FUT.3.SG INT come-NEG-FUT.2.SG INT
 ‘Will he come, won’t you come?’

Complete fusion is not the end of the story. Fusion leads to highly synthetic forms, as a lot of semantic information is centralized in little structural material. As mentioned before, each of these ‘single processing units’ has to be stored individually. In the case of grammatical paradigms where many different morpheme combinations are possible, fusion of each of these to a single processing unit would require a lot of storage capacity. As a “countermeasure”, lesser used forms are “expelled”. This leads to the disappearance of morphologically complex forms, and even of whole inflectional paradigms. These “disappear” gradually, as the first step, in the case of verb form paradigms, is reduction to the third-person. A stylistic correlate of this evolution is the disappearance

from the most-used register, i.e., everyday spoken language, and maintenance in more conservative, i.e., formal and literary, registers. Finally, the whole paradigm falls into oblivion and is replaced by another one, which is typically periphrastic, or at least unmarked compared to the obsolescent forms.

We will next see some examples of morphological erosion in both nominal and verbal paradigms. As can be observed from

(34), which shows different forms of the words for “friend” in some Romance languages, much of the synthetic nominal morphology of Latin has been lost in its daughter languages. The Italian and Spanish examples show that nominative and accusative cases have merged. The corresponding syntactic functions are no more expressed morphologically, but syntactically, i.e., by word order, which was rather free in Latin. It is interesting to see that in Italian plural forms, it is the nominative which has been generalized, whereas in Spanish, the original accusative forms have come to represent both nominative and accusative³. Morphological erosion is most advanced in French: the spoken form of a word like */ami/* says nothing about case, number or even gender, at least outside liaison contexts where final phonemes of older stages may be conserved.

(34) Loss of Grammatical Marking in Nominal Paradigms in Romance

number	gender	case	Latin	Italian	Spanish	French
	m	NOM	amicus	amico	amigo	[ami]
		ACC	amicum			
Sg	f	NOM	amica	amica	amiga	
		ACC	amicam			

³ This presentation is of course a simplification, as it omits all the intermediary stages between classical Latin and the modern Romance varieties. For a more detailed account of the diachronic evolution of nominal morphology in Romance, see Banniard 1997.

	m	NOM	amici	amici	
		ACC	amicos		amigos
Pl	f	NOM	amicae	amiche	
		ACC	amicas		amigas

What about the other cases, such as dative, or genitive? The study of the evolution of the paradigm clearly shows that there are several options when synthetic morphology is lost. We have seen that nominative and accusative are not replaced by analytic constructions, except, maybe, the preposition *a* before certain direct objects in Spanish, as in (35), where the referent of the direct object is emphatic and therefore needs special marking, but not in (36), which contains an “ordinary”, less individuated, object.

(35) Busc-o al profesor
 search-PRS.1.SG DEF.OBJ.SG.M teacher
 ‘I am looking for the teacher.’

(36) Busc-o el nombre de este profesor
 search-PRS.1.SG DEF.SG.M name of D1.SG teacher
 ‘I am looking for the name of this professor.’

When we have a look at the other cases, we see that we have to do with iconicity. This means that morphological marking corresponds to functional markedness. In Spanish, the nominative, being the most central case, is structurally unmarked, and the accusative, which indicates the second argument, is sometimes marked. Dative as well as genitive are always marked, either by affixes as in Latin (37), or by adpositional constructions as in the Spanish example (38).

(37) epistula amic-orum
 letter friend-GEN.PL.M
 ‘the friends’ letter’

- (38) la carta de los amig-os
 DEF.SG.F letter of DEF.PL.M friend-PL.M
 ‘the friends’ letter’

Similar changes where morphology is lost and replaced by alternative constructions can be found in many other languages. Here is another example from German: In conservative registers, the indefinite pronoun *jemand* “somebody” receives suffixes for genitive, dative, and accusative case. In colloquial registers, in contrast, the suffixed versions are disappearing and case relations are expressed, as in the Romance examples, by syntactic position or adpositional constructions.

(39) Loss of Nominal Morphology in German

English	case	“old” paradigm	“new” paradigm
“somebody” (subject)	NOM	<i>jemand</i>	<i>jemand</i>
“somebody’s”	GEN	<i>jemandes</i>	<i>von jemand</i>
“to somebody”	DAT	<i>jemandem</i>	<i>jemand</i>
“somebody”(direct object)	ACC	<i>jemanden</i>	<i>jemand</i>

Getting back to Romance, let’s recall that the examples (23), (24), and

(34) show that French, compared to the neighbouring Romance languages Italian and Spanish, has a particular status with respect to morphological change. The number of inflectional suffixes has been tremendously reduced and the surviving suffixes display heavy syncretism. This means that suffixes aren’t distinctive any more, as one form can refer to different grammatical categories. These suffixes are semantically empty structural residues of what was once a complete paradigm. Inflectional morphology has been conserved much better in Italian, where all of the eight grammatical categories shown in (40) are marked distinctively, whereas in

(spoken) French the number of suffixes for this sample has been reduced to two.

(40) Loss and Syncretism of Verbal Suffixes in French

Grammatical category	Italian	written French	spoken French
PRS.1.SG	<i>cant-o</i>	(je) chant-e	ʃāt
2.SG	<i>cant-i</i>	(tu) chant-es	
3.SG	<i>cant-a</i>	(il) chant-e	
1.PL	<i>cant-iamo</i>	(on) chant-e	
3.PL	<i>cant-ano</i>	(ils) chant-ent	
2.PL	<i>cant-ate</i>	(vous) chant-ez	ʃāt-e
INF	<i>cant-are</i>	chant-er	
PTCP.PF	<i>cant-ato</i>	chant-é	

The fact that once different endings have merged even on the cognitive level becomes obvious when one looks at French writing practice where confusion between the endings *-er*, *-é*, and *-ez* is pervasive. Some readers will object that all the distinctions continue to be expressed in French, namely by the pronominal clitics. This is true, but the obligatoriness of these clitics is only a consequence of the syncretism of the verbal endings.

Syncretism is not the only form of morphological erosion. In other paradigms, the marked members, instead of merging with more frequent forms, fall out of use, and only the unmarked members of the paradigm survive. A good example here is, once again, the French verbal system, which, like the Basque one, has lost many of its synthetic verb forms. Simple past ('*passé simple*') and past subjunctive ('*subjonctif de l'imparfait*') have virtually disappeared from the spoken language, but are maintained in higher registers of the written language. In contrast to other paradigms, however, they are mostly restricted to the third person. A frequency count of inflected forms of the verb *chanter* "to sing" in the FRANTEXT corpus revealed that 95% of the attested past subjunc-

tive forms are third person (*chantât, chantassent*). For simple past (*chanta, chantèrent*), the ratio was similar, namely 93%. Paradigms which are not concerned by obsolescence allow of more variation: the periphrastic present perfect (*a chanté*) is third person in 74% of the cases, and only 55% of future forms (*chantera*) are third person.

German too has synthetic verb forms which have fallen out of use, or are on the way of doing so. This concerns for example past subjunctive ('Konjunktiv II') forms. However, some verbs resist much better to the replacement by periphrasis than others, as illustrated by (41), which is based on a frequency count on the Internet (www.google.com).⁴

(41) Ratio between Present Tense and Past Subjunctive Forms in German

infinitive	A: present 1.sg	B: past subjunc- tive 1.sg	A/B
<i>haben</i> "to have"	324,000	96,500	3.4
<i>sein</i> "to be"	298,000	63,600	4.7
<i>können</i> "can"	181,000	71,300	2.5
<i>dürfen</i> "may"	59,400	3,110	19.1
<i>bringen</i> "to bring"	27,600	400	69.0
<i>riechen</i> "to smell"	4,360	21	207.6
<i>kriechen</i> "to creep"	1,620	2	810.0

We can see that past subjunctive forms of verbs which are also used as auxiliaries are relatively frequent (*hätte, wäre*). In the case of *haben*, the present indicative form (*ich*) *habe* is only 3.4 times more frequent than the corresponding past subjunctive form (*ich*) *hätte*. Similarly, modal verbs are often used in past subjunctive.

⁴ I am aware of the methodological problems of Internet counts, as the validity of the data is often questionable. In the present case, sequences such as pronoun plus inflected verb form, instead of the verb forms alone, helped to reduce the number of invalid responses.

Moreover, since auxiliaries and modals are already part of periphrastic constructions, the periphrastic alternatives to the past subjunctive forms would result in long verbal sequences and are therefore dispreferred. Individual lexical verbs, however, are only used with low frequency, so that the synthetic forms of these verbs cannot be stored as single processing units (Bybee 2001:17-18; Bybee & Hopper 2001:16). Otherwise stated: It would be a waste of storage capacity to treat low frequency synthetic forms of low frequency verbs as autonomous units. Consequently, these forms fall out of use first.

But what happens when certain forms are no longer available? One initial correlate of morphological obsolescence is stylistic shift. This means that the results of once fully productive formations are progressively being avoided in spontaneous speech. They may, however, survive much longer in formal, mostly written, registers where variation is desired and production slower. In the language users' minds, said verb forms then become progressively associated with higher registers, which speeds up their ban from colloquial speech.

Well, this is only about forms, about structure, about signifiers. What happens to the signified, the meaning, the speech intention? The relation between form and meaning is generally a loose one. Forms can have different meanings, a sentence can be interpreted in different ways, and, vice versa, one meaning can be expressed in a lot of different ways. In morphology, this means that speakers will opt for related forms. These are not necessarily periphrastic. Here is another example from Romance: In most Romance languages, past subjunctive forms are used in irrealis conditional clauses (42, from Italian). In French, however, where the past subjunctive is obsolescent, it is the imperfect which is used in conditional subordinated clauses (43).

(42) Se I' av-essi sap-uto
 if it have-PAST.SUBJ.1.SG know-PF.PTCP
 non sarei ven-uto
 Not be:COND.1.SG come-PF.PTCP
 'If I had known it, I wouldn't have come.'

(43) Si je I' av-ais su
 if I it have-IMPF.1.SG know:PF.PTCP
 je ne serais pas ven-u
 I not be:COND.1.SG not come-PF.PTCP
 'If I had known it, I wouldn't have come.'

Interestingly, in Italian itself, the past perfect subjunctive alternates with the imperfect (44), though not with the same temporal reference as in French (45):

(44) Se lo sap-evo non ven-ivo
 If it know-IMPF.1.SG not come-IMPF.1.SG
 'If I had known it, I wouldn't have come.'

(45) Si je le sav-ais, je te le dir-ais
 if I it know-IMPF.1.SG I you:DAT it say-COND.1.SG
 'If I knew it, I'd tell you.'

These examples show that obsolescent synthetic forms can be replaced by other, more vital, synthetic forms. However, periphrastic forms are often preferred because they can be newly constructed when appropriate synthetic forms are not available. Quite often, synthetic and analytic forms co-exist with semantic differences. English displays a variation similar to Basque between synthetic (*goes*) and periphrastic forms (*is going*), but it should be noted that in English it is the periphrastic form which is progressive, whereas in Basque, it is the synthetic form (*doa*). But as we have seen, this

distinction is neutralized in varieties where the synthetic form is not used any more.

5. Conclusion

All the above examples show fusion and replacement, i.e., the creation of (more) synthetic structures on the one hand, and the spread of (more) analytic constructions on the other hand. Traditional explanations for this and similar kinds of variation prefer functional considerations:

- (46) There is a long tradition in functionalist linguistics that attributes a large part of linguistic variation to the interplay of the two opposite motivations of ECONOMY and CLARITY. The cyclic changes are then explained in the following way: for reasons of economy grammatical elements are formally reduced until they are barely recognizable, so that the counteracting motivation of clarity must come in. In order to be understood, speakers then introduce fuller, periphrastic elements, which in turn may become subject to the tendency toward economy. (Haspelmath 1999:1050-1051)

However, these are not the only factors contributing to diachronic change. Moreover, it seems too simplistic to explain different kinds of change with reference to only these two principles. They cannot explain why some forms of a paradigm disappear earlier than others, and how the paradigm is gradually reduced and finally reorganized.

It is therefore worthwhile to examine in more detail the competing motivations underlying the morphological changes under discussion. In this respect, synthetic and analytic inflections are two

strategies used in the construction of forms conveying information on several grammatical categories. It appears that synthesis is more practical for often occurring combinations. You don't have to construct forms each time when you need them, and as you need certain forms quite frequently, you create more compact ready-made forms by morphemic reduction. This is not very practical for less used complex forms. If you create ready-made forms for these as well, you end up with thousands of forms. Where this happens, "natural selection" reduces the stock to the most frequent forms, whereas those which are used less fall into oblivion. For these cases, you are better served with analytic forms you can combine as necessary. The problem is, however, that the frequent and the less frequent forms are part of the same paradigms. Inside the paradigm, analogy is at work, so that the tendency towards synthesis affects all its members. More precisely, morphemic reduction of the frequent forms obscures morpheme boundaries, a process known as 'morphemic obliteration' (Matisoff 1991:385). When the different morphemes lose their formal identity, combining morphemes to yield less frequent forms become difficult. Consider the object prefix⁵ *d-* in the Basque auxiliary form *d-eza-ke-t* 'I can ... it': as long as it was segmentable, there was no problem to put other prefixes in the same position. But what happened then was that, as a result of "routinization"⁶, the sequence *d-eza* was reanalysed as a "single processing unit" *deza*, with the consequence that *eza* was no longer available to form *zait-(e)za-ke-t* 'I can ... you.' Moreover, since the less frequent forms are functionally and structurally marked, they are longer than the frequent forms, which con-

⁵ Etymologically, it would be more correct to analyze this prefix as a tense marker, the third person object being referred to by zero-marking in opposition to the other categories. Synchronically however, the prefix may be considered as an agreement marker, as it occupies the same position as the other agreement prefixes.

⁶ For remarks on 'routinization' c.f., Haspelmath 1999:1055-1062.

tributes to their being avoided by speakers. Results of what was initially a regular formation come to be considered as structurally overloaded:

- (47) The descriptive, iconic function—which in the realm of grammar has been called diagrammatic—is but one of the functions of language in the complex dynamic process of communicating. Sometimes the functional principle “one form: one meaning”, which represents the basis of the diagrammatic strategy, may result in noneconomic bundles of forms, not easy to process and to memorize ... An exceedingly diagrammatic structure may come to represent a growing burden for the (short-term) memory and thus ultimately violate the principle of economy (Ramat 1992: 556-557).

For these cases, a more analytic, periphrastic construction is preferred. Such a construction may have been existing in the language for a long time, but it will come to a new life when selected as an alternative for the collapsing paradigm. The condensed forms survive for a while in the language, but are likely to be replaced once the new paradigm is dominant.

We can sum up by stating that analogy between more and less frequent forms creates pressure to harmonize the paradigm. This harmonization pressure first works to condense all the forms of the old paradigm, but leads to the exclusion of the less frequent of the so created ready-made forms. For these, a new paradigm is emerging, and for a while, the two paradigms co-exist. For the frequent forms, speakers prefer the old synthetic paradigm, for the less frequent combinatorial forms the new analytic paradigm is preferred. The second wave of harmonization integrates the last resisting synthetic forms to the new paradigm. In summary, adaptation to discourse requirements leads to different treatments of frequent and

less frequent forms, whereas the cognitive bias towards harmonization and analogy favours equal treatment. These two antagonistic forces are thus responsible for the constant change in inflectional paradigms.

References

- Altuna, P. & M. Azkarate. 2001. *Euskal Morfologiaren Historia*. Donostia: Elkarlanean.
- Banniard, M. 1997. *Du Latin aux Langues Romanes*. Paris: Nathan.
- Bybee, J. 2003. Mechanisms of Change in Grammaticization: The Role of Frequency. Available at URL <<http://www.unm.edu/~jbybee/mechanisms%20of%20change..doc>>. Published in: Joseph, Brian & Janda, Richard D. 2003 (eds.), *Handbook of Historical Linguistics* 602-623. Oxford: Blackwell (Blackwell Handbooks in Linguistics).
- Bybee, J. & P. Hopper. (eds.). 2001. *Frequency and the Emergence of Linguistic Structure*. Amsterdam & Philadelphia: John Benjamins (*Typological Studies in Language* 45).
- Götz, T. 1991/1995. *Zur Typologie des Sprachwandels*. Heinrich-Heine-Universität Düsseldorf: Magisterarbeit. Available at URL <<http://www-public.rz.uni-duesseldorf.de/~movtchan/magisterarbeit>>.
- Haspelmath, M. 1999. Why is Grammaticalization Irreversible? *Linguistics* 37, 1043-1068.
- Lafitte, P. 1944/2001. *Grammaire Basque (Navarro-Labourdin Littéraire)*. Donostia & Baiona: Elkarlanean.
- Matisoff, J. 1991. *Areal and Universal Dimensions of Grammatization in Lahu* 383-453.
- Ramat, P. 1992. Thoughts on Degrammaticalization. *Linguistics* 30, 549-560.
- Trask, R. 1997. *The History of Basque*. London & New York: Routledge.
- Trask, R. 1998. The Typological Position of Basque: Then and Now. *Language Sciences* 20.3, 313-324.
- Traugott, E. & B. Heine. 1991. (eds.). *Approaches to Grammaticalization*. Amsterdam: Benjamins.
- Zubiri, E. & I. Zubiri. 2000. *Euskal Gramatika Osoa*. Bilb(a)o: Didaktiker.

Corpus

<http://www.egunkaria.com> (new address: www.berria.info)
http://www.euskaracorpUSA.net/XXmendea/Konts_arrunta_fr.html
<http://www.herriak.info>
<http://zeus.inalf.cnrs.fr/noncateg.htm> (FRANTEXT)

Grammatical category labels

ABL	Ablative	LOC	Locative
ABS	Absolutive	MR	Modifier
ALL	Allative	NEG	Negation
COND	Conditional	NR	Nominalizer
D	Deictic	PAST	Past
DAT	Dative	PF	Perfect
DEF	Definite	PL	Plural
DET	Determiner	POT	Potential
ERG	Ergative	PROG	Progressive
FUT	Futur	PRS	Present
HYP	Hypothetical	PTCP	Participle
IMPF	Imperfect	REL	Relative
INSTR	Instrumental	SG	Singular
IO	Indirect object	SUBJ	Subjunctive