# From polyfusional to post-fusional: Obsolescence and innovation in Basque predicate morphosyntax and its typological implications

Gerd Jendraschek
University of Regensburg[*]

## Abstract

Basque possesses a complex verbal morphology, where the inflected verb, i.e. the predicate core, shows agreement with (or represents) up to three arguments, namely ergative, absolutive, and dative marked noun phrases. This entails that the syntactic relations between the core arguments and the verb complex are perfectly reflected in verbal and nominal morphology, by both head- and dependent marking. Moreover, Basque predicates indicate tense (present, past, hypothetical), mood (indicative vs. subjunctive) and realis vs. irrealis. This leads to a large number of possible forms, with a high degree of morpholological irregularity. While the replacement of finite lexical verbs by complex predicates involving auxiliaries is well documented, the gradual obsolescence of many auxiliary forms is an ongoing and less noted phenomenon. The Basque agreement pattern illustrates the syntactic hierarchy between the core arguments in different argument constellations. The argument at the top of the syntactic hierarchy, i.e. the subject, allows of most variation. Head-marking cross-referencing arguments further down on the hierarchy is increasingly restricted, so that in some cases only third person dependents can be cross-referenced. Basque thus exhibits both the *Monotransitive* and the *Ditransitive Person-Role Constraint*, albeit to different degrees. Auxiliaries corresponding to uncommon argument constellations fall gradually out of use. Frequency is therefore not only responsible for grammaticalization, but also for the subsequent disintegration of paradigms according to the *Rarity Condition on Obsolescence*. As a consequence of the obsolescence of verbal morphology, new strategies for the expression of syntactic relations (and semantic content) emerge. The evolution illustrated in this paper is a symptom of the typological shift from agglutinative to fusional morphology and beyond, and suggests that a very high degree of synthesis is incompatible with a high degree of fusion. In addition to language-internal factors, the influence of Romance contributes to the obsolescence as well as the preservation of morphologically complex forms.

*Keywords:* Basque, Historical Linguistics, Obsolescence, Morphological Typology, Auxiliaries

# 1 Introduction

Basque verbal morphology is a prime example of language change in progress. A well-known phenomenon (cf. Trask 1998:319) is the gradual replacement of simple predicates, where all inflectional markers are directly attached to the lexical verb, by complex predicates, where most of those markers are hosted by an auxiliary verb which combines with an aspectually marked participle of the lexical verb. Example (1) illustrates the synthetic inflection of the verb root *kus* 'to see' (Altuna & Azkarate 2001:152). This form is not used any more. (2) shows the periphrastic inflection which has replaced it.

(1)  *Dakusat.*
    da    -kus    -at
    PRS  -see    -A.1SG[1]
    'I see (it).'

(2)  *Ikusten dut.*
    ikus -ten    d    -u    -t
    see -IMPF  PRS  -AUX  -A.1SG
    'I see (it)'

To reconcile recognizability of forms and analytical depth, examples are presented in a four-line arrangement which better suits languages with a high degree of morphological complexity: The first line shows the unanalyzed object language example with punctuation and capitalization as appropriate, the second line shows morpheme boundaries, the third line presents interlinear morphemic translations (more on this in §2), and the last line provides a free translation.

In the 16th century, synthetic forms of the type *dakusat* in (1) were attested for about sixty verbs (Lafon 1943, quoted in Trask 1998:319 and Haase 1992:87). This number has gone down to between ten and twenty verbs, with the exact number dependent on sociolinguistic factors. To a large extent, this diachronic change is recent, as can be seen when comparing, for example, the morphology of Northern Basque as described in Lafitte's grammar (1944/2001) with the morphological performance of young speakers of Northern Basque today. Many of the forms listed in Lafitte's grammar, which described the Basque language as it was used in the first half of the 20th century, have simply disappeared within less than a century. In Southern Basque, the loss has been slower, so that forms that are no longer used and often unknown in the North are still used in the South, albeit with stylistic connotations, as they are considered literary or formal (Jendraschek 2003a:40−41).

Whether all verbs could be inflected synthetically at some historical stage is an open question which goes beyond the scope of this paper. Haase (1992:87) argues that periphrastic inflection was a contact-induced phenomenon to accommodate loan verbs

---

[1] The following abbreviations are used in this paper: A, transitive subject; ACC, accusative; ALL, allative; AUX, auxiliary (root); BEN, benefactive; COND, conditional; DAT, dative; DET, determinative; ERG, ergative; FUT, future; GEN, genitive; HYP, hypothetical; IND, indicative; IO, indirect object; IMPF, imperfective; NEG, negative; O, (direct) object; PRF, perfective; PL, plural; PNCT, punctual; POT, potential; PROG, progressive; PRS, present; PST, past; RL, realis; RFL, reflexive; S, intransitive subject; SG, singular; SUBJ, subjunctive.

from Latin, and from there, spread to the indigenous stock.

In present-day Basque as illustrated in (2), the root of the verb 'to see' is *ikus*. As a consequence of the diachronic change from the synthetic inflection in (1) to the more analytic inflection in (2), not only have the cross-reference and tense affixes changed their host, moving from the root *kus* to an 'auxiliary root' *u*, also the initial *i* in *ikus* has fused with the original root *kus*, so that *ikus* must now be analyzed as a single morpheme. The same is true for verbs which still have synthetic forms, such as *ibili* 'to walk; to function'. The citation form, which is identical to the perfective participle, is *ibili*, where the initial *i* has to be considered part of the root *ibil*; it has no function in synchrony. The synthetic forms, however, are *nabil* 'I walk', *dabil* 's/he walks' etc. Here, we could isolate a root *bil*. The loss of synthetic inflection therefore has as indirect consequence the disappearance of a morpheme boundary, as the initial slot, that is the *i* on *kus* and *bil*, becomes invariable. The remaining synthetic forms have to be stored individually. Compare, for example, first person singular *nabil* 'I walk' with second person singular *zabiltza* 'you walk'. The final element *tza* has no synchronic status. That it used to denote plural reference (as one of many plural allomorphs) is an irrelevant reconstruction as the second person plural form of *ibili* is now *zabiltzate* (where *te* can be analyzed as denoting plurality). Morphological fossils such as *i* and *tza* make it impossible to inflect verbs according to patterns which were historically regular but no longer are. Such loss of morphological motivatedness and transparency is called 'morphemic obliteration' in Matisoff (1991:385) and 'demorphemicization' in Lehmann (2002:12).

Another example of morphemic obliteration are causatives formed with *ra*. The causative of *ikusi* 'to see' is *erakutsi* 'to show'. We can reconstruct the pre-root prefix *e~i*, the causativizer *ra*, the root *ku(t)s*, plus the perfective aspect suffix *i* used for citation forms. Synchronically, however, such segmentation would not make any sense, as this pattern is not productive any more.

Initially the reduction of the inflectional possibilities of lexical verbs simplified verbal morphology. Rather than inflecting the sixty or so synthetic verbs (or possibly at some historical stage the whole verb inventory of Basque) separately, one could just use the same auxiliaries with all kinds of verbs. While this diachronic process has been well described, it is not the end of the story, as the same process which affected the lexical verbs is now affecting the auxiliaries. (3) shows the archaic synthetic inflection, with a potential morpheme *ke* directly attached to the root *tor* 'to come'. (4) shows the current periphrastic form in Southern Basque. The root of the lexical verb is now *etor*, and the potential suffix is incorporated into the auxiliary verb (where the gloss AUX represents the verb root). For the morphological make-up of the auxiliaries, see §2 and §3. (5) shows the last stage: the bound potential morpheme *ke* has been replaced by a free particle *ahal*, thereby reducing the number of inflected auxiliary forms. This construction is dominant in Northern Basque.

(3)　　*Datorke.*
　　　　da　　　　　-tor　-ke
　　　　PRS(S.3SG)　-come　-POT
　　　　'S/he can/may/will come.'

(4)  *Etor daiteke.*
     etor    da          -i      -teke
     come   PRS(S.3SG)   -AUX   -POT
     'S/he can come.'

(5)  *Etortzen ahal da.*
     etor   -tzen   ahal   d          -a
     come  -IMPF   POT    PRS(S.3SG)   -AUX
     'S/he can come.'

This marginalization of morphological forms, first *datorke*, then *(etor) daiteke*, is an example of morphological obsolescence, as described e.g. in Dorian (1978), where the term refers to 'simplification in morphological performance' and 'movement toward simplicity in a highly complex morphology' (1978:606). However, while Dorian's examples are taken from a dying language, the Basque language as a whole is very well alive, albeit in certain places and contexts it remains in a diglossic situation with Romance languages. If we add the increasing importance of English, Basque has been and continues to be in intense contact with 'Standard Average European'. The sociolinguistic situation of Basque is indeed very dynamic. Old domains are lost, while new ones are conquered, both trends happening at the same time. At the end of this paper, I will try to address the question to what extent the sociolinguistic situation impacts language change.

At this stage, it is important to note that **structural obsolescence** is not limited to communities that undergo **sociolinguistic obsolescence**: the gradual erosion of an entire language system correlating with language shift is just an extreme case of something that also happens in vibrant monolingual communities (the Basque case being in between these extremes). Certain parts of the grammar (and the lexicon) may become obsolete, while innovative grammatical (and lexical) phenomena emerge and spread. The interaction between morphological obsolescence and grammatical innovation will also be addressed in this paper.

One may wonder how the alternative strategies emerge in the first place. At least for the alternatives shown in examples (1) and (2), and to some extent (3) and (4), we know that the synthetic and periphrastic constructions originally had, and for certain verbs still have, different functions. For those verbs that synchronically preserve both synthetic and periphrastic inflections, the periphrastic form expresses habitual aspect, whereas the simple form has a progressive reading (Zubiri & Zubiri 2000:412–413). The synthetic form *dator* (i.e. (3) without the potentiality suffix) is therefore more accurately translated as 'is coming (now)' whereas *etortzen da* corresponds to 'comes (regularly)'. What started as a synchronic variation to express semantic distinctions then becomes a diachronic change as the construction involving fewer inflectional forms, which in this case yields the analytic predicate *etortzen da*, expands into new domains. As the semantic distinction becomes blurred, the old form is considered redundant, becomes a mere stylistic variant, and eventually falls into oblivion.

A similar change has occurred in colloquial German, where most preterite forms (*wusch* 'washed'), once semantically distinct from the perfect forms (*hat gewaschen* 'has washed'), have been replaced by the latter, so that the aspectual difference has been neutralized (Dudenredaktion 2009:514). The loss of preterite forms is attested since the 15th century, and is more advanced in the spoken language and in the southern dialects, but has

been spreading to the written standard as well as northern colloquial varieties of Standard German (cf. Schrodt & Donhauser 2003:2518). This shows again that whole grammatical subsystems can disappear, in this case the preterit forms of German, without any external cause such as language contact or shift.

While the synthetically inflected forms of lexical verbs in Basque (*dakusat* 'I see', *nabil* 'I walk', *datorke* 's/he can come') are uniformly replaced by periphrastic inflections (*ikusten dut, ibiltzen naiz, etor daiteke*), the response to the obsolescence of the auxiliaries (e.g. *dut, naiz, daiteke*) is not so uniform. This paper is the first attempt at a systematic overview of avoidance and replacement strategies, including loss of semantic content and agreement, case shift involving the downgrading of complements to adjuncts, non-finite constructions, lexical paraphrases, etc. While specific changes have been described for specific dialects (e.g. Arregi & Nevins 2007 and 2008; Hualde 2000 and 2002), this paper focuses on standardized varieties, showing that obsolescence continues despite standardization. In fact, as Basque is promoted from a predominantly oral variety to an official language, diachronic processes are reinterpreted as synchronic layering. The alternative strategies will be discussed in §4. To facilitate the understanding of that section, §2 will give a short introduction to the argument-verb agreement pattern of Basque, followed by some theoretical prerequisites in §3. Finally, §5 concludes that morphological obsolescence, beginning at the periphery of the system, helps to bring down the extremely high indices of synthesis and fusion. In other words, the incompatibility of polysynthesis and fusion is transforming Basque from a 'polyfusional' into a 'post-fusional' language.


## 2  The Basque Agreement Pattern

§2 is intended as a brief introduction to the Basque agreement pattern, beginning with monovalent and ending with trivalent constructions. In present-day Basque, most verbs are inflected periphrastically. This means that the lexical verb appears as a participle marked for aspect (perfective, imperfective, prospective, and neutral); the auxiliary it combines with indicates the argument structure (cross-reference with absolutive, ergative, and dative-marked NPs), tense (present, past, hypothetical), modality (realis vs. irrealis/potential), and mood (indicative, subjunctive, imperative). Auxiliary roots are suppletive, and may be zero as a result of phonological erosion. In this paper, I have opted for 'etymological glossing', essentially based on the analyses in Trask (1997) and Hualde (2003b), i.e. a segmentation policy that disregards diachronic obliteration of morpheme boundaries. This has the advantage of better showing the internal complexity and evolution of auxiliaries, but is often speculative with regards to morpheme boundaries and meanings. Another downside is that it does not reflect how Basque speakers handle inflection, as the analysis here is often based more on historical reconstruction than on synchronic paradigmaticity. Synchronically, many of the forms discussed below, or recurrent parts of them, would be better characterized as fossilized cumulative morphemes.


### 2.1  Agreement with a single argument

When the verb has a single argument, the intransitive subject (S) is cross-referenced or represented by prefixes (3rd person being implicit in the tense marker, see below). When the

predicate is periphrastic, as in (6) to (8), these prefixes are hosted by the auxiliary. Auxiliaries have suppletive roots depending on valency (intransitive vs. transitive) and mood (indicative vs. subjunctive). From now on, these parameters will be indicated in the glosses for the auxiliary roots (instead of a generic AUX), provided the root is segmentally identifiable. Some of the cross-reference markers originated as cliticized forms of personal pronouns, such as *ni* '1SG', *zu* '2(PL>SG)', or *gu* '1PL' (cf. Trask 1997:218). Second singular forms like *zara* are historically (and therefore morphologically) plural forms. The present tense morph *da* has the allomorphs *d* (2) and *a* (6).

(6)　*Iristen naiz.*
　　　iris　-ten　　n　　-a　-iz
　　　arrive -IMPF　S.1SG　-PRS -IND.INTR
　　　'I arrive.'

(7)　*Iristen zara.*
　　　iris　-ten　　z　-a　-r　-a
　　　arrive -IMPF　S.2　-PRS -PL　-IND.INTR
　　　'You arrive.'

(8)　*Iristen gara.*
　　　iris　-ten　　g　　-a　-r　-a
　　　arrive -IMPF　S.1PL　-PRS -PL　-IND.INTR
　　　'We arrive.'

## 2.2  Agreement with two arguments

Table 1 summarizes and examples (9) to (12) illustrate agreement with an absolutive and an ergative argument on present-tense auxiliaries. Here, the pre-root slots reflect the absolutive-marked O argument, whereas the ergative-marked A is cross-referenced by suffixes.

|  |  | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
|  |  | O | tense | number of O (old system) | root | number of O (innovative) | A | number of A |
| SG | 1 | *n* | *a* |  | *u* |  | *t* |  |
| SG | 2 | *z* | *a* | *it* | *u* |  | *zu* |  |
| SG | 3 | *Ø* | *d* |  | *u* |  | *Ø* |  |
|  |  |  |  |  |  |  |  |  |
| PL | 1 | *g* | *a* | *it* | *u* |  | *gu* |  |
| PL | 2 | *z* | *a* | *it* | *u* | *zte* | *zu* | *e* |
| PL | 3 | *Ø* | *d* | *it* | *u* |  | *Ø* | *(z)te* |

Table 1. Present-tense auxiliaries with two arguments

Third person O is not marked for person, but for number if plural, and is inferrable because of its distinctive tense allomorph *d-* (*-a* elsewhere).

(9)  *Ikusten dut.*
ikus -ten    Ø    -d   -u        -t
see  -IMPF   O.3  -PRS -IND.TR-A.1SG
'I see him/her/it.'

(10)  *Ikusten duzu.*
ikus -ten    Ø    -d   -u        -zu
see  -IMPF   O.3  -PRS -IND.TR-A.2
'You see him/her/it.'

(11)  *Ikusten ditut.*
ikus -ten   Ø   -d  -it   -u       -t
see  -IMPF  O.3 -PRS -O.PL  -IND.TR-A.1SG
'I see them.'

(12)  *Ikusten nauzu.*
ikus -ten      n    -a  -u       -zu
see  -IMPF    O.1SG -PRS -IND.TR-A.2
'You see me.'

There is also a bivalent intransitive agreement pattern, where the predicate cross-references an S in absolute case and a dative-marked indirect object.

## 2.3 Agreement with three arguments

When the clause contains an (overt or recoverable) indirect object in addition to the two central arguments, the corresponding cross-reference marker in the auxiliary is inserted between the markers for O and A. The different markers are summarized in Table 2.

|      |   | (1) (O) | 2 tense | 3 IO | 4 number of O | 5 IO | 6 IO | 7 A |
|------|---|---------|---------|------|----------------|------|------|-----|
|      | 1 |         |         |      |                |      | *-da/-t* | *t* |
| SG   | 2 |         |         |      |                |      | *-zu* | *zu* |
|      | 3 |         | *d*     | *i*  |                |      | *-o* | Ø |
|      | 1 |         |         |      |                |      | *-gu* | *gu* |
| PL   | 2 |         |         |      |                |      | *-zue* | *zu* |
|      | 3 |         | *d*     | *i*  | *z*            | *ki* | *-e* | Ø |

Table 2. Present-tense auxiliaries with three arguments

The marker *i~ki* in slots 3 and 5 seems to have originated as a valency-increasing device for

adding an indirect object (Trask 1997:228). It will be glossed IO for lack of a better term. So positions 3 (if O is singular) and 5 (if O is plural) in Table 2 indicate that the following cross-reference marker in position 6 agrees with or represents the dative argument. This explains the term 'dative flag' which is sometimes used for *(k)i* (e.g. Trask 1997:227). We would expect the root *u* after position 3, but it must have been lost in indicative mood (Trask 1997:234); for an example involving the subjunctive root *eza*, see (29).

Here as before, the arguments need not be overtly expressed, as the cross-reference markers alone can represent participants. Consequently, the brackets [] in (13) indicate that the clause is grammatically well-formed with or without these elements.

(13) *[Nik zuri liburuak] ematen dizkizut.*
    ni  -k      zu   -ri      liburu -a          -k
    1SG -ERG    2    -DAT     book   -DET(ABS)   -PL

    ema -ten       Ø   -d  -i   -z    -ki  -zu  -t
    give -IMPF     O.3 -PRS -IO  -O.PL -IO  -2   -A.1SG
    'I give them [the books] to you.'

(14) *Ematen dizkidazu.*
    ema -ten       Ø   -d  -i   -z   -ki  -da   -zu
    give -IMPF     O.3 -PRS -IO  -PL  -IO  -1SG  -A.2
    'You give them to me.'

Table 2 reveals an important syntactic restriction. When the clause contains an indirect object, the direct object can only be third person. This is shown by leaving position 1 in Table 2 empty. We will see in §4.2 what happens in ditransitive constructions with direct objects having first- and second-person referents.


# 3  Morphological Erosion in the Emerging Standard

## 3.1  Variation in Basque dialects and in the standard(s)

The introduction to Basque verbal morphology in §2 has only shown a small fraction of possible forms. In the rest of the paper, we will see that morphological complexity increases further if we vary the tense and (certain types of) modality. Moreover, the illustration in §2 was limited to forms from Standard Basque. If we take into account dialectal forms, the variety of patterns found becomes even more astonishing, as certain dialects display morphological changes that are absent from the standard (see e.g. Hualde 2002 on the levelling of 'ergative displacement' in Bizkaian).

Nevertheless, looking at the relatively young standard language is very instructive from a sociolinguistic point of view. The emerging standard is clearly a compromise between an abstract, one might say fictitious, deregionalized Basque – promoted under the label *euskara batua* 'unified Basque' – on the one hand, and various dialectal 'substrates' on the other (Jendraschek 2007:110–116). The standardization has thereby transformed a 'family' of localized, to some extent mutually unintelligible (cf. Hualde 2003a:3) varieties into a common, yet pluricentric language (cf. Clyne (ed.) 1992). This means that Basque

speakers do not use just one, but several standards. These can be roughly described as Northern, South-Western, and South-Eastern Basque, corresponding to the three political entities of the Northern Basque Country (in France), the Basque Autonomous Community, and Navarre (in Spain). These standards are themselves in a contact situation with each other. The morphological erosion that has been under way for centuries in the dialects continues in Standard Basque, but with different implications. The regional and diachronic variation has certainly been reduced, while the multiplicity of competing forms and constructions has been partially reinterpreted as a stylistic (or 'diaphasic') layering (see the division into five 'groups' in Table 6). At the same time, a construction that is common, even in a colloquial register, in South-Western Basque may be perceived as conservative and formal (or even be unknown) in Northern Basque (but less commonly the other way round).

## 3.2 Agglutinative vs. fusional morphology: a useful idealization

The observation that standard Basque perpetuates the typological evolution that has been under way for centuries in the dialects is indicative of functional pressures underlying this variation. The simultaneous expression of the categories tense, (certain types of) modality, and cross-reference with up to three arguments in a single inflected verb form produces a large number of possible forms. However, in all categories some terms are clearly underrepresented, which is an indicator of their markedness. There is often an iconic correlation between low frequency (due to markedness in discourse) and morphological marking. This means that functionally marked forms are also structurally more complex. In contrast, more frequent (functionally unmarked) terms also tend to be structurally unmarked, i.e. represented by shorter or even zero morphemes (lack of morphological encoding within a paradigm, as opposed to overt marking for other values of the same paradigm).

The constellation of structural complexity and low frequency marginalizes marked forms. As a consequence, they fall out of use. The most frequent forms, in contrast, can be stored as 'single processing units' (cf. Bybee 2001:619–622; Bybee & Hopper 2001:16; see also Haspelmath & Sims 2010:272–273, and for a similar argumentation with respect to change in Bizkaian, Hualde 2002). At a later stage, those morphologically complex forms which will not be stored as such, due to low frequency, cannot be constructed regularly *either* after the paradigm has 'shifted' from an agglutinative to a fusional type (as morpheme boundaries have become opaque). Uncommon forms disappear first from colloquial speech (less affected by normative considerations and at the same time requiring faster processing), but may survive in higher registers (more on this 'register gap' in §5.3).

Such an explanation complements the 'Frequency Condition on Grammaticalization' (Haspelmath 2004:27), which states that frequent expressions are more likely to undergo grammaticalization, with the frequency itself being pragmatically and semantically motivated. The phenomenon described here might accordingly be labelled 'The Rarity Condition on Obsolescence', which states that less frequent expressions are more likely to succumb to obsolescence. The two 'conditions' could be combined to a 'Frequency Condition on Coalescence', which states that frequent expressions first undergo morphologization and then demorphemicization (cf. Lehmann 2002:12; 132–137; Song 2005:795), whereas the rare expressions either do not undergo morphologization in the first

place, or if they do, later fail to demorphemicize and are 'left behind' on the grammaticalization path where they are marginalized as the remnants of a defunct paradigm, and eventually proceed directly to loss, the last stage of coalescence.

This explanation is an attempt to fit morphological obsolescence into the grammaticalization model, but it equates demorphemicization with morphological obsolescence, which is a bit of a stretch. First, demorphemicization corresponds to the transition from agglutinative to fusional morphology (Lehmann 2002:12). However, for morphological obsolescence it is of secondary importance whether the *obsolescent* forms are agglutinating or fusional at this stage. What is essential is that the *resilient* forms are becoming fusional, reduce their internal complexity, and are eventually stored as single processing units, which makes them independent from their original paradigm. Second, loss as a consequence of phonological attrition is once again what may happen to the *frequent* forms, whereas the loss of *obsolescent* forms is a consequence of the breakdown of an inflectional arrangement. The corresponding paradigms disintegrate, but their remaining members undergo fossilization rather than attrition. We will have a chance to continue the discussion about the role of synthesis and fusion in language change in §5.2.

The distribution of values in Table 3 reflects the general direction of obsolescence in Basque (cf. Jendraschek 2003a:31–42), and may be different for other languages (and even certain paradigms in Basque itself, cf. Hualde 2002). The functional pressure eliminates marked values first, before spreading to forms representing unmarked values (Arregi & Nevins 2007 have a similar list of marked environments, 'in which impoverishment is likely to occur').

| marked values | unmarked values |
| --- | --- |
| ------------------------------------> | |
| past | present |
| oblique | subject |
| first/second person | third person |
| plural | singular |
| irrealis/potential | realis |

Table 3. Direction of morphological erosion

While the distinction between 'agglutinating' and 'fusional' is more of a useful idealization than an empirically verifiable fact (see e.g. Haspelmath 2009), the comparison between inflection in Basque and Turkish shows that the idea behind this distinction is typologically relevant, as it has potential implications for the synchrony and the diachrony of paradigm organization, which, as we will see, is quite different in Turkish and Basque. In a fully operative system (e.g. in an 'ideal' agglutinative language, Turkish being a language that comes close to that ideal), all categories can be combined as necessary: morphemes are overwhelmingly non-cumulative, morpheme boundaries are unobscured, fusion phenomena are limited to regular phonological processes such as vowel or consonant assimilation, allomorphy is relatively predictable, and the order of morphological slots shows little variation.

These conditions hold for Turkish, but not for Basque. In (15) and (16), we compare two inflected Turkish forms. If we postulate zero-morphemes for morphologically unexpressed values (representing semantically relevant information), both forms have five

morphological slots.

(15) *Onu bana veriyorsun.*
on -u    ban -a    ver -Ø  -iyor  -Ø  -sun
3SG -ACC   1SG -ALL    give -RL  -PROG  -PRS -2SG
'You are giving it to me.'

(16) *Onu bana verebildin.*
on -u    ban -a    ver -ebil-Ø    -di -n
3SG -ACC   1SG -ALL    give -POT-PNCT   -PST -2SG
'You could give it to me.'

The order of morphemes is the same in the two forms:
   1. verb stem *ver*;
   2. POTENTIAL *ebil* vs. Ø for REALIS;
   3. PROGRESSIVE aspect *iyor* vs. Ø for PUNCTUAL;
   4. PAST tense *di* vs. Ø for PRESENT tense;
   5. person-number, here 2nd singular *(su)n.*

Even Turkish is not perfectly agglutinative though. The 1st person singular pronoun *ben* has an allomorph *ban* when followed by the allative case suffix, and the third person pronoun *o* has the allomorph *on* before any case suffix. The second person singular cross-reference suffix is either *sun* (we ignore predictable vowel harmony here) or *n* depending on the type of overt morpheme that precedes it (in our examples aspect (15) vs. tense (16)).

     Compared to Basque however these idiosyncracies appear negligible. Unlike Turkish, where the second and third argument of trivalent verbs are expressed by free pronouns (which because of that freedom do not interact with verbal inflection), inflected forms in Basque have bound, and to some degree fused, pronominal marking. As pointed out before, the glossing provided for (17) and (18) is an attempt to segment fossilized forms.

(17) *Ematen didazu.*
ema -ten    Ø    -d  -i  -da -Ø  -zu
give -IMPF    O.3  -PRS -IO  -1SG -RL  -A.2
'You give it to me.'

(18) *Eman zeniezadakeen.*
eman  z  -en -i  -eza    -Ø -da    -ke -en
give   A.2 -PST -IO  - SUBJ.TR  -O.3 -IO.1SG -POT -PST
'You could (=were able to) give it to me.'

Furthermore, the different inflectional categories interact with each other. Whereas the indicative auxiliary can combine with aspectual forms of the lexical verb (here imperfective), the subjunctive auxiliary combines with an aspectually neutral form. The subjunctive itself is triggered by the potential morpheme *ke* (at least in contemporary Southern Standard Basque). This makes it difficult to assign the semantic value of possibility to just one morpheme, comparable to *ebil* in the Turkish example (16). Moreover, there is no indicative morpheme in (17). The suppletive root *u* seen in the bivalent

transitive forms (9)-(12) would have been lost in forms like *didazu*, as a consequence of phonological reduction (Trask 1997:234).

In the present tense, the cross-reference marking for the A argument is the last of the cross-reference markers (17), in the past tense example (18) it is the first. When a 'suffix', it has the form *zu*; when a 'prefix', it has the form *z*, but is difficult to segment from the following past tense morpheme *en*, which in this paradigm appears only if the A is morphologically 1st or 2nd plural. An alternative segmentation for the cross-reference marker would therefore be *zen*, a paradigm-specific allomorph of *z*. I put 'suffix' and 'prefix' in quotation marks as I would argue that a segmentation of these forms into a root plus affixes (as attempted here) is an etymological exercise with no synchronic justification. This explains why a learner-oriented grammar such as Zubiri & Zubiri (2000:461) only divide the form into the four morphemes *zenieza-da-ke-en*.

The tense marking in (18) illustrates well the problem to identify and correctly parse homonymous segmental strings. A suffix with the form *(e)n* is e.g. also found in subjunctive forms with optative meaning (*ikus deza-**n**!* 'may he see') or relative clause forms (*ematen didazu-**n** oparia* 'the present which you are giving me'), and can therefore not be unambiguously associated with past tense. And, while past and hypothetical have different markers when the subject is third person (*eman **z**-iezadakeen* 's/he **was** able to give me' vs. *eman **l**-iezadake* 's/he **would be** able to give me'), for other subjects the only difference is the absence of final *en* in the hypothetical (*eman zeniezadake-Ø* 'you would be able to give me').

As hinted at in §2.2, the tense/mood-markers *d* for present tense, *z* for past, and *l* for hypothetical indicate at the same time third person (of O, S, or A, depending on the actual form), for they fill the same initial slot as the cross-reference markers in 1st and 2nd person. Here, person-marking interacts with tense-marking, and *d/z/l* are in fact often presented as 3rd person allomorphs whose distribution depends on tense. It is therefore difficult to associate morphological units with semantic units, which is the hallmark of an agglutinative system. As a consequence, these forms are difficult for speakers to process (and for linguists to gloss) and cannot be constructed, but, like irregular inflectional forms elsewhere, are fossilized remnants of a system that was undoubtedly more transparent in a distant past. Similar to the Gaelic variety described in Dorian (1978:591), Basque inflects words in a 'rich variety of ways', and like there, this 'richness is essentially gratuitous'. In both languages, it is therefore prone to succumb to morphological simplification (Dorian 1978:606).

The absence of regularities helping to parse and construct inflected forms as required by discourse has consequences for the less frequent forms. The form *didazu (17)* appears 226 times in euskaracorpusa (an online corpus of Basque 20th century literature) and I would say that all fluent speakers can use it without problems, but *zeniezadakeen (18)* appears only once in the same corpus (in a dialogue from a stage play). In Turkish, a form like *verebildin* is in no way obsolescent, but a full member of an intact paradigm; every native speaker of Turkish would understand and use such a form. In Basque, a form like *zeniezadakeen* is in contrast classified as archaic and rare (e.g. in Hualde 2003b:230), so that some otherwise fluent speakers may not even understand such a form (Jendraschek 2003a:41).

It should have become obvious that Turkish and Basque morphology function very differently. Much of Basque verbal morphology is not 'productive' anymore, by which I mean that the forms are to a large extent fossilized, which in turn favours the emergence

of alternative forms and patterns, as will be illustrated in §4. In Turkish, the high frequency patterns support low frequency forms, as the same construction rules apply. In Basque, in contrast, the different paradigms are morphologically too different to support each other. The high-frequency forms and paradigms then become autonomous within the system, thereby exposing the remaining forms to obsolescence. Thus, while morphological complexity is not inherently difficult for speakers, its survival presupposes either a high functional load of the complex forms, or at the least indirect support from analogous high-frequency forms on which the low-frequency forms can then be modelled. It is therefore not the semantic constellation of marked values *per se* that leads to the obsolescence of the forms expressing it, but the heterogeneity of inflectional principles ('fusional' vs. 'agglutinating', in our idealization) that leads to the marginalization of certain forms.[2]

Further evidence for the breakdown of older verbal morphology comes from some Basque dialects, where speakers have abandoned irregular patterns and prefer more regular innovative forms. In the Bermeo variety of Bizkaian, past tense forms are often modelled on the present tense forms: whereas both standard Basque and literary Bizkaian show distinct patterns for present and past tense formation as illustrated in Table 4, the Bermeo dialect has given up most of these old forms and replaced them with more transparent, albeit still synthetic, forms (Hualde 2002).

| meaning | standard Basque | lit. Bizkaian | Bermeo dialect |
|---|---|---|---|
| 'we V it' (present; A.1PL; O3SG) | *dugu* | *dogu* | *du* |
| 'we V-ed it' (past; A.1PL; O.3SG) | *genuen* | *genduen* | *dun* |
| 'I have brought it to him' (present perfect; A.1SG; O.3SG; IO.3SG) | *ekarri diot* | *ekarri deutsat* | *ekarritzat* |
| 'I had brought it to him' (past perfect; A.1SG; O.3SG; IO.3SG) | *ekarri nion* | *ekarri neutzan* | *ekarritzaten (~ekarri nozan)* |

Table 4. Innovative past tense forms in Bermeo Basque (adapted from Hualde 2002)

In both examples (*du→ dun*, *ekarritzat → ekarritzate**n***) we can see that Bermeo Basque basically simply adds a past tense morpheme *(e)n* to the present tense form, where both standard Basque and literary Bizkaian have a different order of morpheme slots. Interestingly, Bermeo Basque has a few traces left of the older forms, such as the form *nozan* 'I had V-ed it to him', clearly related to the corresponding form in literary Bizkaian. Unlike the remaining argument constellations, where the older forms have been lost, this form competes with the new innovative form. The reason why it has not yet been completely 'evicted' from the paradigm is likely to be 'the greater frequency of use of this form' (Hualde 2002).

The Bermeo data are interesting for being a rare case of entirely new synthetic forms, whereas the innovative strategies shown in §4 are either periphrastic or extend the

---

2 Although not the topic of this paper, let me mention that Basque and Turkish also differ in their nominal morphology. In Turkish, number and case markers do not fuse; in Basque they do (ak marks both ergative singular and absolutive plural, *ek* marking ergative plural).

use of existing forms into new domains. Before proceeding to a discussion of these patterns, I will attempt to explain the typological motivation that triggers them in the first place.


## 3.3 Central vs. peripheral morphological forms

I have argued that in a language like Turkish, which comes close to the agglutinative ideal, the expression of combinations of marked values is not a problem in terms of morphological productivity. Basque however is a long way away from the agglutinative ideal – an already complex system has been further opacified by morphological obsolescence. As another example of how markedness is involved in morphological obsolescence in Basque, consider (19) to (23). The direct object of transitive constructions can be freely exchanged in realis modality (see (19) and (20)), although here too, third person accounts for about 95% in (admittedly, written) corpora[3], but not in potential modality (21 vs. 22).

(19)  *Ikusten dut.*
  ikus -ten      Ø     -d    -u          -t
  see  -IMPF   O.3   -PRS -IND.TR -A.1SG
  'I see him/her/it.'

(20)  *Ikusten zaitut.*
  ikus -ten      z   -a   -it     -u          -t
  see  -IMPF   O.2  -PRS -O.PL  -IND.TR -A.1SG
  'I see you.'

(21)  *Ikus dezaket.*
  ikus   Ø    -d   -eza       -ke    -t
  see    O.3  -PRS - SUBJ.TR  -POT   -A.1SG
  'I can see him/her/it.'

(22)  *%Ikus zaitzaket.*
  ikus z   -a  -it    -za       -ke    -t
  see  O.2 -PRS -O.PL   -SUBJ.TR  -POT   -A.1SG
  'I can see you.'

(22) represents marked values for both person and modality. It is still used by some speakers (as indicated by '%'), but many never use this form. The more a form represents marked values, the higher its proneness to obsolescence. This explains why a form like *zintzakedan* 'I could … you' in (23) is considered archaic by speakers to such an extent (if they understand the form at all) that it is virtually inexistent even in corpora of written Basque (and classified as 'very rare' in Hualde 2003b:229).

---

[3] Written corpora of Basque include articles collected from newspapers (*egunkaria* and *berria*), the literary corpus *euskaracorpusa*, edited by the Royal Academy of the Basque Language and the Basque Centre for Terminology and Lexicography, as well as Internet counts.

(23) *Ikus zintzakedan.*
    ikus   z   -in  -tza      -ke   -da   -n
    see    O.2 -PST - SUBJ.TR -POT   -A.1SG -PST
    'I could see you.'

Examples (19) to (23) show that basic oppositions can still be expressed in some paradigms, but not in others. Moreover, the obsolescence of (22) and (23) illustrates the *Monotransitive Person-Role Constraint*, the 'preference for Agents to be first/second person and Patients to be third person' (Haspelmath 2004:50). The crucial difference is (a) that this may not affect all paradigms with cross-reference to a Patient argument, but only the less frequent forms (cf. the resilient counter-example *zaitut* in (20)); and (b) that these constructions are marginal not because they would not have 'become grammaticalized' (Haspelmath 2004:51), but, on the contrary, because they are ejected from a grammaticalized system of which they had once been fully integrated members (cf. Hualde 2003b:229). This is further evidence that the Frequency Condition on Grammaticalization and the Rarity Condition on Obsolescence complement each other.

      Before we proceed, some further observations on the morphological make-up of the auxiliaries. The indicative realis (=non-potential) counterpart of (23) is given in (24).

(24) *Ikusi zintudan.*
    ikus -i     z   -in   -t   -u    -da     -n
    see -PRF   O.2   -PST   -O.PL  -IND.TR-A.1SG -PST
    'I saw you.'

In this paradigm, the root *u*, which we also saw in *dut* (19) or *zaitut* (20), alternates with *du*, which occurs in the forms for historically singular direct objects, as in (25).

(25) *Ikusi ninduzun.*
    ikus -i     n     -in -du   -zu -n
    see -PRF   O.1SG -PST -IND.TR-A.2 -PST
    'You saw me.'

Hualde (2003b:222) suggests that the *t* in forms like *zintudan* is a remnant of the plural marker *it*, with the sequence *in-it-du* 'PST-O.PL-AUX' subsequently reduced to *intu*. Note the misleading surface similarity between *zintzakedan* (23), *zintudan* (24), and *zintezke* (26). They all contain an initial string *zint*, which, however, has to be glossed differently in the three forms.

(26) *Joan zintezke.*
    joan   z  -in -te    -z   -ke
    go     S.2 -PST - POT   -S.PL-POT
    'You could go'

Similar morphological 'homonymy' can be seen in (20) and (54), where the sequence *zait* is made up of entirely different morphemes. Such lack of unique morpheme-meaning associations further complicates the system.

      If we try to reconstruct the underlying morphemes of the auxiliary forms in (19)-

(24), we can isolate the segments listed in Table 5. The fact that the current second person singular forms were historically plural forms (marked by *(i)t*) that expanded and were subsequently restricted to singular addressees (Trask 2003:151) further illustrates the discrepancy between such a reconstruction and a purely synchronic description of the form.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| A | tense | number of O | root/mood/ valency | reality | A | tense |
| Ø | d~a | Ø | u | Ø | t~da | n |
| O.3 | PRS | O.SG | IND.TR | RL | A.1SG | PST |
| z | in | (i)t | (e)za | ke | | Ø |
| O.2 | PST | O.PL | SUBJ.TR | POT | | NON-PST |

Table 5. Underlying morphological slots

The fact that the historical glossing is a reconstruction that does not correspond to the cognitive reality of the speakers underlies the obsolescence of the forms with second-person O. For discursive reasons, the sequence *d-eza* (21) occurred much more often than the sequence *z-a-it-za* (22)(third person O just being more common). It is plausible to postulate a state where the system was less fusional so that all the units listed in Table 5 were separately available morphemes. Speakers could take them from their morpheme inventory and combine them as necessary. Morphemic obliteration and routinization led to the reanalysis of *deza* as a single morpheme, which became part of the speakers' morpheme inventory. As a consequence, *eza* disappeared from the morpheme inventory. As for the sequence *zaitza*, it was much less frequent than *deza*, and therefore never entered the morpheme inventory in its fused form. With neither *zaitza* nor *eza* being part of the inventory of frequently used morphemes, forms like *zaitzaket* 'I V you' were marginalized, and pushed to the periphery of the linguistic system. This reanalysis of agglutinative forms as fusional forms led to the reorganization of verbal morphology from a closed inventory of short non-cumulative morphemes to a hierarchically organized system of fused cumulative forms.

More should be said about the distinction between the centre vs. the periphery of a linguistic system. The discrepancy between the combinational potential of verbal morphology on the one hand, and the frequency of forms in real-language corpora on the other has important consequences for an adequate description of the language as well as its development. The possible number of forms is immense. This is why we find large numbers of tables with paradigms in most grammatical descriptions of Basque. However, these descriptions produce a false impression, as present-day spoken Basque uses only a small subset of all possible forms. The two approaches – listing all the possible forms vs. listing frequently attested forms – can be unified though, if we locate possible forms along a cline leading from the centre of the grammatical system to its periphery (a similar approach can be found in King 1993:190; see also Haase 1992:159–174). Whereas the centre contains the most frequent and stable forms, the outer periphery contains forms which exist only in theory. Such a peripheral form would not violate any grammatical rule; however, speakers will avoid it, as their primary aim is not to exploit the morphological potential of the

system, but to make themselves understood. In Jendraschek (2007:192–198), the possible forms of auxiliaries in potential mood (i.e. those containing the suffix *-ke*) were divided into five groups, from the most central to the most peripheral: frequent, medium-frequency, conservative, very rare, and obsolete. The characterization of the forms is based on a) frequency counts in online corpora (online newspapers, EUSKARACORPUSA, GOOGLE), b) observations of systematic (diachronic, diatopic, diaphasic) variation, and c) speaker judgements on some of the forms. Below, I will briefly summarize the five groups.

a. The FREQUENT FORMS appear in all written texts, including those written in northern Basque where such forms are obsolescent in the spoken language. For a majority of speakers, however, these forms are very frequent even in the spoken language. The form *dezaket* (21) 'I can V (it)' is in this group.

b. The MEDIUM-FREQUENCY FORMS are less common in colloquial speech, except for 'conservative' speakers, as they give the speech a more elaborate flavour. In contrast, they are very important in the written language.

c. The CONSERVATIVE FORMS are characteristic of high registers and can be used to enrich a text or in specific contexts. Many speakers have only passive knowledge of these forms

d. The VERY RARE FORMS are avoided by most speakers. Their use may even lead to communication problems as their internal complexity makes them difficult to process. Where they appear in texts, they seem unnatural and are often the result of a too literal translation. The forms *zaitzaket* (22) 'I can V you' and *zintzakedan* (23) 'I could V you' are in this group.

e. The OBSOLETE FORMS are usually only found in grammar books. They are generally absent from texts written in modern Basque.

A schematization of this system is presented in Table 6.

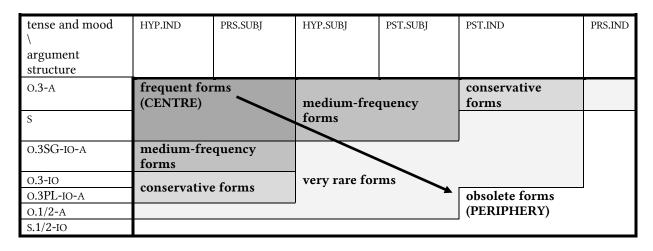| tense and mood \ argument structure | HYP.IND | PRS.SUBJ | HYP.SUBJ | PST.SUBJ | PST.IND | PRS.IND |
|---|---|---|---|---|---|---|
| O.3-A | **frequent forms (CENTRE)** | | **medium-frequency forms** | | **conservative forms** | |
| S | | | | | | |
| O.3SG-IO-A | **medium-frequency forms** | | **very rare forms** | | | |
| O.3-IO | **conservative forms** | | | | | |
| O.3PL-IO-A | | | | | **obsolete forms (PERIPHERY)** | |
| O.1/2-A | | | | | | |
| S.1/2-IO | | | | | | |

Table 6. Centre vs. periphery of the system of auxiliaries in potential modality

It is not necessary here to present all the actual forms; the point is that a combination of an 'unmarked' inflectional value with a 'marked' value already results in a peripheral form. When for example a third person direct object is cross-referenced on a bivalent hypothetical indicative or present subjunctive predicate (yielding forms such as *ikusiko luke* 's/he would see' or *ikus dezake* 's/he can see'), the result is a frequent form from the centre of the grammatical system. In contrast, when one tries to construct an extended intransitive clause (ABS-DAT) with a non-third person subject, the tense-mood categories remaining the same, the resulting forms *etorriko nintzaioke* 'I would come to him/her' *etor nakioke* 'I can come to him/her' are obsolete, i.e. they are clearly located on the periphery of the system. This distribution is schematized in Table 7, which presents the forms for constellations of the most central and the most peripheral values.

| 1ˢᵗ parameter ╲ 2ⁿᵈ parameter | HYP.IND (unmarked) | PRS.IND (marked) |
|---|---|---|
| O.3SG-A (unmarked) | *luke* (CENTRE) | *duke* (PERIPHERY) |
| S.1SG-IO (marked) | *nintzaioke* (PERIPHERY) | *natzaioke* (PERIPHERY) |

Table 7. Combination of marked and unmarked values

Let me add a note on the present indicative forms. Their location on the periphery might come as a surprise, as it seems to be in contradiction to the markedness cline shown in Table 3, where indicative forms were presented as less marked, and therefore more central. The explanation is that the modal morpheme *ke*, due to its inherent semantic value, combines more easily with non-actual categories, such as hypothetical (yielding forms equivalent to 'would') and subjunctive (yielding forms equivalent to 'can'), whereas the combination of *ke* with present and past indicative had a low functional load, and was therefore relegated to the periphery of the system. While Table 6 only presents a subdomain of inflected auxiliaries (namely those containing the modal marker *ke*), the concepts of centre vs. periphery can be applied to other grammatical systems where an inflation of morphological forms leads to a hierarchical organization of the system. The more complex a system is, the more it is bound to become internally structured. One pattern of internal organization is to rank the members by establishing a hierarchy, with some members in the centre, others on the periphery. The periphery has fringes with fuzzy boundaries, so that it is not always clear whether a peripheral form should actually (still) be counted as a member of the system. For the expression of some of the constellations on the periphery of the system, alternative strategies develop. This is what happens in Basque predicate morphosyntax, and will be the topic of §4.

## 4  Syntactic Strategies in Response to Morphological Obsolescence

§3 has shown how auxiliary forms corresponding to uncommon value-constellations fall gradually out of use, and as one example we have seen how this erosion in verbal morphology makes it impossible for first- and second-person participants to appear as arguments in certain constellations. As a response to morphological obsolescence, Basque

speakers have developed several alternative strategies for the expression of uncommon participants.


## 4.1 Redistribution or loss of semantic content

One way to avoid complex forms is to transfer components of their meaning to separate words, i.e. by using more analytic constructions. The auxiliary *zaitzaket* in (27) expresses agreement, tense, and modality, whereas the form *zaitut* in (28) does still express agreement and tense, but not potential modality, which is then expressed by a separate particle (*ahal* in (28))[4].

(27)　*% Ikus zaitzaket.*
　　　ikus z　-a　-it　　-za　　　-ke　　-t
　　　see　O.2　-PRS　-O.PL　-SUBJ.TR　-POT　　-A.1SG
　　　'I can see you.'

(28)　*Ikusten [ahal] zaitut.*
　　　ikus -ten　　ahal z　-a　-it　　-u　　　　-t
　　　see　-IMPF　POT　O.2　-PRS　-O.PL　　-IND.TR-A.1SG
　　　'I [can] see you.'

This decentralization of information (i.e. by increasing analyticity) can simplify the morphology of the auxiliary considerably. In (29) vs. (30), we have a difference of two morphemes, as the meanings of SUBJ and POT are transferred from the auxiliary to the particle *ahal*.

(29)　*Eman diezazkizuket.*
　　　eman Ø　-d　-i　-eza　　　-z　　　-ki　-zu　-ke　-t
　　　give　O.3　-PRS　-IO　-SUBJ.TR　-O.PL　-IO　-2　　-POT -A.1SG
　　　'I can give them to you.'

(30)　*Ematen [ahal] dizkizut.*
　　　ema-ten　　　ahal Ø　-d　-i　　-z　-ki　-zu　-t
　　　give-IMPF　　POT　　O.3　-PRS　-IO　-O.PL-IO　-2　　-A.1SG
　　　'I [can] give them to you.'

The periphrastic *ahal*-construction is widespread in northern Basque, but its use is limited in southern Basque, where semantic loss, i.e. neutralization of the modal opposition between realis vs. potential, is a more common way to avoid complex auxiliaries. An alternative construction without semantic loss is embedding, which will be shown in §4.3.

---

[4] In classical Basque, possibility was expressed by a combination of the particle *ahal* and the modal suffix *ke*. Modern Basque varieties usually choose one or the other, but it is clear that the particle construction is morphologically simpler. For a detailed study of the expression of possibility in Basque, see Jendraschek (2007), for a shorter description Jendraschek (2003b).

Still another option is to switch to a similar, but a more morphologically familiar modality type, using conditional instead of possibility forms, as in (31). Such an 'avoidance strategy' is originally a pragmatic choice, but it may lead to language change if it is conventionalized.

(31)   *Emango nizkizuke.*
       eman-go   n      -i   -z     -ki -zu -ke
       give-PROSP A.1SG -IO -O.PL -IO -2   -POT
       'I would give them to you.'

A radical yet frequent solution is to leave syntactic relations unexpressed when they can be inferred from context, as is often the case with beneficiaries and recipients. This reduction strategy is particularly frequent in requests (King 1993:337). If a speaker pronounces (32) while standing in front of a closed door, it is unnecessary to specify that they will be the beneficiary of the door's being opened. The explicit (33) is therefore considered as (too) formal or as a calque from Romance.

(32)   *Atea ireki zenezake?*
       ate -a     ireki  z   -en -eza     -ke
       door-DET   open  A.2 -PST -SUBJ.TR -POT
       'Could you open the door (for me)?'

(33)   *Atea ireki zeniezadake?*
       ate -a     ireki  z   -en -i  -eza     -da -ke
       door-DET   open  A.2 -PST -IO  -SUBJ.TR -1SG -POT
       'Could you open the door for me?'

Use of the citation form (which corresponds to the perfective participle and also functions as imperative) is not considered impolite, just informal (and therefore the usual construction in most oral contexts). The lack of auxiliaries in this construction, which is certainly one reason for its popularity in present-day colloquial Basque, allows for the optional insertion of pronouns, like *guri* 'to us' in (34). In more formal contexts, (35) may be found, where the auxiliary form *iezaguzu* encodes the argument-constellation.

(34)   *Bidali [guri] mezu bat edo deitu [guri].*
       bidal -i  gu -ri     mezu    bat edo dei -tu    gu  -ri
       send   -PRF 1PL -DAT   message one or   call -PFV   1PL -DAT
       'Send [us] a message or call [us].'

(35)   *Bidal iezaguzu mezu bat edo dei iezaguzu.*
       bidal  Ø  -Ø  -i  -eza     -gu -zu mezu    bat
       send   O.3 -IMP -IO  -SUBJ.TR -1PL -A.2 message one

       edo dei Ø   -Ø  -i  -eza     -gu -zu
       or   call O.3 -IMP -IO  -SUBJ.TR -1PL -A.2
        'Send us a message or call us.'

*Deitu guri!* 'call us!' can be seen as a compromise between the simple (and most genuinely Basque) *deitu!* 'call!' (once again, the unexpressed referent is clear from the context) and the formal *dei iezaguzu*, which, rendered into English, comes close to something like 'may you call us'.


## 4.2 Lack of agreement

We have seen that the auxiliaries agree not only with the arguments in ergative and absolutive case, but also with dative-marked indirect objects. However, it is not possible to have inflected forms that agree with all three arguments if the absolutive argument is first or second person (Laka 1993:27). This is an instance of the *Ditransitive Person-Role Constraint* (Haspelmath 2004). The third participant, i.e. the indirect object, is then left to nominal morphology (the case suffix *-ei* in (37)), as its presence or absence has no influence on the form of the auxiliary, which is *nau* in both (36) and (37).

(36)  *Jonek aurkeztu nau.*
  Jon -ek  aurkez -tu  n  -a -u   -Ø
  Jon -ERG present -PFV O.1SG -PRS -IND.TR-A.3SG
  'Jon has introduced me.'

(37)  *Jonek bere lagunei aurkeztu nau.*
  Jon -ek  bere     lagun -ei
  Jon -ERG POSS.RFL.3SG friend -DAT.PL

  aurkez -tu  n  -a -u   -Ø
  present -PFV O.1SG -PRS -IND.TR-A.3SG
  'Jon has introduced me to his friends.'

Compare now (36) and (37) with (38) and (39), where the direct object is third person, so that agreement with three arguments is not only possible, but even obligatory (but see (40) for colloquial language). In that case, the presence of an (overt or recoverable) indirect object in (39) triggers agreement, so that the auxiliary is *die* instead of *du*.

(38)  *Jonek Xabier aurkeztu du.*
  Jon -ek  Xabier aurkez -tu Ø -d -u   -Ø
  Jon -ERG Xabier present -PRF O.3 -PRS - IND.TR -A.3SG
  'Jon has introduced Xabier.'

(39)  *Jonek Xabier aurkeztu die.*
  Jon -ek  Xabier aurkez -tu Ø -d -i -e -Ø
  Jon -ERG Xabier present -PFV O.3 -PRS - IO -3PL -A.3SG
  'Jon has introduced Xabier to them.'

It is not clear whether lack of agreement in cases like (37) is due to loss of corresponding forms (see Laka 1993:27, Trask 1997:220–221, and Hualde 2003b:223 for remarks on some attested counter-examples) or if there is under-differentiation due to low frequency of the

category (as described by Haspelmath & Sims 2010:268–270). Haspelmath (2004) further suggests that such disharmonic associations of persons (1st/2nd) and roles (Theme) do not undergo grammaticalization in the first place, and are therefore not integrated into the emerging paradigms of bound pronominal markers. Addis (1993:431) describes this phenomenon as 'paradigm trimming' due to 'marking-overload'.

In colloquial speech, we can often find constructions, where the presence of a third argument is irrelevant for inflection (40), presumably resulting from influence of Romance. Remember that such constructions are ungrammatical according to prescriptive grammar.

(40)　*Jonek Xabier haiei aurkeztu du.*
　　　Jon -ek　　Xabier　hai -ei　　aurkez -tu　Ø　-d　-u　　-Ø
　　　Jon -ERG　Xabier　D3 -DAT.PL　present -PRF　O.3　-PRS - IND.TR　-A.3SG
　　　'Jon has introduced Xabier to them.'

For speakers with little exposure to written, formal Basque (where prescriptive grammar is more important), the distinction between two- and three-participant auxiliaries is not clear. This might be a contact-induced confusion, as person markers (clitics and pronouns) in French and Spanish are often the same for direct and indirect objects (see Jendraschek 2006:155). Those speakers then choose the less marked paradigm, yielding *nau* in (42) instead of the required three-participant form *dit* in (43). This phenomenon is also known as dative displacement (see e.g. Fernández 2001)

(41)　*Ekarri nau.*
　　　ekarr -i　　　n　-a　-u　　　-Ø
　　　bring -PRF　　O.1SG -PRS -IND.TR-A.3SG
　　　'He brought me (here).'

(42)　*Ogia ekarri nau.*
　　　ogi　-a　　　ekarr -i　n　-a　-u　　　-Ø
　　　bread -DET(ABS)bring -PVF　O.1SG -PRS -IND.TR-A.3SG
　　　'He brought me the bread.'

(43)　*Ogia ekarri dit.*
　　　ogi　-a　　　ekarr -i　Ø　-d　-i　-t　　-Ø
　　　bread -DET(ABS)bring -PFV　O.3　-PRS -IO　-1SG　-A.3SG
　　　'He brought me the bread.'

Prescriptive grammarians are of course appalled by this evolution. Lafitte (1944/2001:296) remarks:

> Le solécisme de la côte. – Les Labourdins de la côte … prennent le complément direct d'objet au sens de complément-datif. C'est une grosse erreur. … Ces formes sont curieuses, mais il y a lieu de les écarter comme contraires au courant général de la langue basque.
> [The solecism of the coast. – Lapurdians from the coast use the direct object complement instead of the dative complement. That's a big mistake. These forms are curious, but they should be avoided for being contrary to the general tendency of the Basque language.].

Interestingly, in parts of Gipuzkoa (the province whose capital is Donostia-San Sebastian),

there seems to be a mixed paradigm, where this confusion is limited to first-person singular indirect objects.[5] This can be seen in (44.a), which is widespread in those varieties (compare with the 'correct' form *didate* in (44.b)), in contrast to (45), where the standard form *dizute* is maintained, as illustrated by (45.b).

(44)  a.   *Esan naute.*
           esan  n     -a    -u        -te
           say   O.1SG  -PRS  -IND.TR   -A.3PL
           'They told [it to] me.'

      b.   *Esan didate.*
           esan  Ø   -d   -i  -da    -te
           say   O.3  -PRS -IO  -1SG  -A.3PL
           'They told [it to] me.'

(45)  a.   *\*Esan zaituzte.*
           esan  z    -a  -it    -u        -zte
           say   O.2  -PRS -O.PL  -IND.TR   -A.3PL
           'They told [it to] you.'

      b.   *Esan dizute.*
           esan  Ø   -d   -i  -zu  -te
           say   O.3  -PRS -IO  -2   -A.3PL
           'They told [it to] you.'


### 4.3 Non-finite constructions

In some cases, speakers have the possibility to use embedded non-finite constructions where relevant participants appear as overt constituents, as in (47), when the corresponding auxiliaries (*zaitzaket* in 46) are no longer in use.

(46)  *Ikus zaitzaket.*
      ikus  z   -a  -it   -za       -ke   -t
      see   O.2  -PRS -O.PL  -SUBJ.TR  -POT  -A.1SG
      'I can see you.'

(47)  *Zu ikusteko gai naiz.*
      zu        ikus -te  -ko   gai    n     -a  -iz
      2(ABS)    see  -NR  -MR   able   S.1SG  -PRS -IND.INTR
      'I am able to see you.'

---

As opposed to the *ahal*-construction in §4.1, use of the *gai*-periphrasis emphasizes the protagonist's ability to do something, and for that reason it is not semantically equivalent to the potential forms of auxiliaries. This is more or less comparable to the alternation in English between *can/could* and semantically more specific periphrases like *to be able to.* If ability is not emphasized, neutralization of the modal opposition or shifting to more familiar types of irrealis modality (future, conditional) is common (see §4.1).

## 4.4  Demotion to adnominal status

First- and second-person participants can be structurally represented as third-person arguments by demotion to adnominal status. Thus, instead of being a dependent of the verb complex, situation participants appear as possessive pronouns modifying the dummy noun *buru* 'head'. This construction is obligatory for, but also limited to, reflexive situations, as in (48).

(48)   *Nire burua aurkeztu nahi dizuet.*
      ni  -re    buru-a    aurkez-tu    nahi  Ø  -d  -i  -zu  -e  -t
      1SG  -GEN   head-DET   present-PRF   want  O.3  -PRS  -IO  -2  -PL  -A.1SG
      'I want to introduce myself to you.'

Here again (see 49), colloquial speech may have Romance-type agreement where only the subject (and implicitly the direct object) is cross-referenced, whereas the indirect object is left to nominal morphology, in (49) the pronominal form *zuei* 'to you'.

(49)   *Nire burua zuei aurkeztu nahi dut.*
      ni  -re      buru  -a    zu  -ei
      1SG  -GEN   head  -DET   2   -DAT.PL

      aurkez  -tu  nahi  Ø  -d  -u      -t
      present  -PFV  want  O.3  -PRS  -IND.TR-A.1SG
      'I want to introduce myself to you.'

## 4.5  Case shift

In order to avoid complex auxiliaries, speakers may use noun cases which do not require agreement, such as the benefactive in (51) or the allative in (53) instead of the dative shown in (50) and (52).

(50)   *Piarresi ez diozu fitxik ekarri.*
      Piarres  -i      ez  Ø  -d  -i  -o  -zu     fitxik     ekarr-i
      Piarres  -DAT   NEG  O.3  -PRS  -IO  -3SG -A.2    nothing  bring-PRF
      'You brought nothing to Piarres.'

(51) *Piarresendako ez diozu fitxik ekarri.*
Piarres -endako ez Ø -d -u -zu fitxik ekarr-i
Piarres -BEN NEG O.3 -PRS -IND.TR-A.2 nothing bring-PVF
'You brought nothing for Piarres.'

(52) *% [Niri] etortzen bazatzaizkit…*
ni -ri etor -tzen ba -z -a -tza -i -z -ki -t
1SG -DAT come -IMPF COND -S.2 -PRS -IND.INTR -IO -S.PL -IO -1SG
'If you come to me….'

(53) *Niregana etortzen bazara…*
ni -regana etor -tzen ba -z -a -r -a
1SG -ALL come -IMPF COND - S.2 -PRS -PL -IND.INTR
'If you come to me.'

Case shift is only possible in situations which allow for a referent to be encoded with more than one semantic role, e.g. recipient (corresponding to dative case) and beneficiary, one of which is highlighted when selected for structural representation. An alternative which is similar to case shift is lexical paraphrase which is exemplified in §4.6.


## 4.6 Lexical paraphrases

In many cases, simple auxiliary constructions are widespread with third-person cross-reference (as in example (54)), but unusual (or, at least, limited to conservative and literary registers) in the case of first- or second-person participants. Instead of (55) then, speakers may use lexical paraphrases like the one in (56).

(54) *Lagun bat etorri zait.*
lagun bat etorr -i Ø -Ø -za -i -t
friend one come -PFV S.3 -PRS -IND.INTR -IO -1SG
'A friend has come to me.'

(55) *Etortzen bazatzaizkit…*
etor -tzen ba -z -a -tza -i -z -ki -t
come -IMPF COND -S.2 -PRS -IND.INTR -IO -S.PL -IO -1SG
'If you come to me…'

(56) *Nire etxera etortzen bazara…*
ni -re etxe -ra etor -tzen ba -z -a -r -a
1SG -GEN house -ALL come -IMPF COND - S.2 -PRS -PL -IND.INTR
'If you come to my house…'

# 5 Conclusions

§4 has shown how morphological obsolescence is compensated for by alternative constructions. In the process, syntactic relations may (a) remain the same, (b) change while keeping semantic relations relatively constant, or (c) both syntactic and semantic relations change. These different scenarios will be synthesized in §5.1. More generally, the morphological obsolescence we observe in Basque is part of a broader typological shift that started with agglutination, which then evolved into fusional, or even *polyfusional* morphology, which is ultimately being replaced by *post-fusional* analytic morphosyntax. This topic will be explored in §5.2. Although morphological obsolescence in Basque can be explained to a large extent by internal factors, namely extreme degrees of synthesis and fusion, external factors such as language contact and diglossia, do play a role as well, as we will see in §5.3.

## 5.1 The evolution of syntactic relations

§4 has shown different responses to the changes affecting the expression of grammatical relations between the predicate and its arguments in Basque. These alternative strategies help avoid complex auxiliaries that have succumbed to morphological obsolescence. This can be achieved by simplifying the utterance so that less grammatical information needs to be encoded (§4.1), simplification of agreement requirements, either by disregarding certain requirements (§4.2), or by using less central noun cases which do not require agreement (§4.5), or by demoting participants to embedded non-finite constructions (§4.3) or adnominal status (§4.4 and §4.6).

The use of such avoidance strategies has consequences for syntactic relations. The theoretical problem may be stated as follows:

Assuming for every change

(a) a synchronic stage *A*, in which syntactic relation *R* is encoded by a morphological marker *M*; and

(b) a subsequent synchronic stage *A'*, where *M* becomes obsolescent

we describe what happens to *R* in *A'*. The data presented in §4 have illustrated the following three possibilities:

A. *R* remains constant.

This is what happens in the case of loss of encoding, as illustrated in §4.2. It starts as a mismatch between *R* and *M*, but as the syntactic relation is not represented solely by *M*, the syntactic status of the dependent is not affected. This was the case in (42), of which we see an altered version in (57), where the free pronoun *niri* has been added.

(57) *Ogia niri ekarri nau.*
     ogi -a         ni -ri    ekarr -i    n    -a -u        -Ø
     bread -DET(ABS)1SG -DAT  bring -PVF   O.1SG -PRS -IND.TR-A.3SG
     'He brought me the bread.'

The syntactic relation is one of ditransitive verb to indirect object, while *M* consists of

head-marking cross-referencing a direct object complement. The overt NP now representing the cross-referenced complement in (56) reveals that $R$ remains constant as evidenced by dative case marking on the dependent. In this case, $M$ is reanalyzed as cross-reference to an oblique complement rather than an (in)direct object.

B. $R$ changes, semantic relation remains constant.

In many cases, there is a development from head-marking to dependent-marking, as in the replacement of the finite predicate *dei iezaguzu* (stage $A$) by the non-finite *deitu guri* (stage $A'$), both meaning 'call us' (ex. 34-35).
Strictly speaking, $R$ remains constant here too, but not quite: the addition of the free pronoun *guri* to a finite verb form on which it is already cross-referenced would be pragmatically marked, as $M$ is already expressed on the predicate; in the non-finite construction in contrast $R$ is established by $M$. This entails that the syntactic relation is less grammaticalized in $A'$ than in $A$.
Another relevant example is the demotion of a core participant to adnominal status, as seen in §4.4. I repeat the relevant example here for convenience.

(58)  *Nire burua aurkeztu nahi dizuet.*
      ni  -re   buru-a    aurkez -tu   nahi  Ø  -d  -i  -zu  -e  -t
      1SG -GEN  head-DET  present -PRF  want  O.3 -PRS -IO  -2   -PL  -A.1SG
      'I want to introduce myself to you.'

Here, $R$ would represent the underlying mapping of semantic onto syntactic roles. We have the trivalent verb *aurkeztu* 'introduce, present' with the reflexive constellation Agent = Theme, plus a Recipient. The Theme of the trivalent verb in $R$ would canonically correspond to a direct object. However, in the surface realization (= $R'$), the verb complex does not directly govern the Theme; instead it governs the relational noun *buru* 'head' which in turn governs the Theme participant as its possessive complement. From $R$ to $R'$, we thus move from immediate to mediate government.

C. $R$ changes, semantic relation changes.

Here, speakers choose a different semantic representation, but their motivation is syntactic, as they avoid a syntactic relation whose morphological encoding is obsolescent. A relevant example is the choice of benefactive or allative instead of dative case marking as seen in §4.5. While $R$ constitutes government, $R'$ is a case of modification by an adjunct. The complement and the adjuncts not only differ in their syntactic status, but also represent different semantic roles.
In summary, we have seen that the massive loss of morphological options could be compensated for by new strategies, in most cases the renovation through analyticization. We could therefore say that we have functional continuity despite morphological obsolescence, provided we adopt a relatively broad conception of function, namely one that is derived from the communicative goals. However, if we look at semantic relations and their syntactic representations in more detail, we find that both may change. The changes described for Basque are thus a good illustration of the interplay between morphology, syntax, semantics, and pragmatics.

## 5.2 Typological shift

The evolution illustrated by the data of §4 is the consequence of a general loss of morphological transparency. Traditionally, languages with overall high morphological transparency, i.e. with a rather high degree of synthesis but a low degree of fusion, have been called 'agglutinative', those with low morphological transparency 'fusional'. Hence, a more general perspective on morphological typology may be in place here. When distinguishing between 'agglutinative' vs. 'fusional' morphology, several criteria can be used. Table 8 shows those listed in Plank (1999:282–284):

| Morphological markers can be |
| --- |
| 1. separatist or cumulative |
| 2. invariant or variant |
| 3. distinct or homonymous |
| 4. always or never or sporadically zero |
| 5. local or extended |
| 6. repeatable or once-only |
| 7. members of larger or smaller paradigms |
| 8. transparently or opaquely segmentable |
| 9. loosely or tightly cohesive |
| 10. loosely or tightly bound |
| 11. optional or obligatory |

Table 8. Parameters relating to morphological fusion

For the characterization of verbal morphology as a whole (rather than individual markers), we also need to take into account the **index of synthesis**, i.e. the 'number of morphemes per word' (Comrie 1989:46). And while all the criteria in Table 8 can be considered independently, it seems reasonable to conceive segmentability, i.e. the **index of fusion**, as a combination of various criteria from Table 8. The high degree of synthesis has led to qualifying Basque as an agglutinative language, see e.g. Echenique (1997:152):

> Es sabido que la lengua vasca es tipológicamente una lengua aglutinante, por lo que las referencias a los objetos (directo, indirecto) aparecen en la misma forma verbal ... Como lengua aglutinante, el vasco integra en el verbo, por ejemplo, las marcas de ergativo, acusativo (sic!) y dativo, junto a los morfemas y lexemas verbales.
> [It is known that the Basque language is typologically an agglutinative language, whence references to (direct, indirect) objects appear in the same verb form. As an agglutinative language, Basque integrates in the verb, for example, ergative, accusative, and dative markers, together with verbal morphemes and lexemes.]

However, the high index of fusion shows that inflectional morphology in Basque is not agglutinating, as evidenced by the large number of irregular and to a large extent fossilized paradigms, including many cases of suppletion. The two parameters and the typical language types they yield are summarized in Table 9.

| index of synthesis<br>index of fusion | *low* | *high* | *extreme* |
|---|---|---|---|
| *high* | n/a | fusional | polyfusional? |
| *low* | isolating | agglutinating | polysynthetic |

Table 9. Morphological typology

Table 9 tentatively introduces the term *polyfusional* for languages that simultaneously have an extreme degree of synthesis and a high degree of fusion, contrasting with *polysynthetic* languages, which have a lower ratio of fusion to synthesis. Comrie (1989:49) surmises that 'as the index of synthesis gets higher, the ratio of agglutination to fusion must also increase' as 'it is inevitable that a language with a very high index of synthesis will also have a low index of fusion' (p. 46). We would therefore expect that in a language, or a subsystem within that language, the diachronic development of extreme degrees of both synthesis and fusion will eventually lead to the collapse of fusional inflection.[6] This is confirmed by the Basque scenario described in this paper. The first step has been almost completed: the elimination of synthetic forms of *lexical* verbs, such as *dakusat* 'I see it' in (1) and *datorke* 's/he may come' in (3). The second step is ongoing: the elimination of complex, multiply marked, and therefore peripheral forms of *auxiliary* verbs, as seen in §4.

What we have seen in Basque is the conflict between the requirement to express complex semantic content on the predicate on the one hand, and the lack of transparent morpho-semantic associations on the other. Learners of Basque very soon notice a contradiction between the inflectional 'potential' (the sum of all the forms allowed by the system), and what they hear in actual discourse. Grammar books, and therefore Basque language education both for first and second language speakers, tend to reflect the potentially available forms. Obviously, all these forms are grammatically correct, but that does not mean that they are pragmatically adequate; in fact, many of those forms are unknown to ordinary speakers. As illustrated by Table 6, they belong to the periphery of the system, and only a minority of all possible forms can be considered central, with forms in between being distributed according to sociolinguistic parameters such as diatopic, diaphasic, and diastratic variation.

The collapse of morphological paradigms is therefore gradual: Basque morphology is still working well with those auxiliaries that are less affected by morphological obsolescence (see also Jendraschek 2003a:31–42), but for the less frequent of the synthetic forms, the more analytic devices shown in §4 become dominant. The irregularities brought about by the shift towards a higher degree of fusion make certain verb phrases cognitively difficult to handle. These are relegated to the periphery of the system, and eventually new strategies for the expression of syntactic relations (and semantic content) emerge. The old morphological system is reduced to the core of the system, which keeps shrinking. Such a system may be called *post-fusional.* The present paper has described this evolution for Basque. Contrary to Echenique's claim, Basque is thus not an agglutinative language at all, as this would presuppose a much lower degree of fusion, but a polyfusional language that is becoming post-fusional.

---

[6] Though a strong sociolinguistic cohesiveness of the community may help to preserve irregular patterns; see Croft (2000:192–193) and Jendraschek (2007:326–338) on the relationship between the external and the internal history of a language.

The incompatibility of polysynthesis and fusion is one factor likely to trigger morphological obsolescence. This is not an automatic and necessary evolution. Obsolescence can be caused by other factors as well, such as sociolinguistic disruption and language contact (see e.g. Dorian 1978 and §5.3). On the other hand, some polysynthetic languages do have elements of fusion, such as Greenlandic and Chiricahua Apache (Aikhenvald 2007:8). Classical Basque is another promising candidate for a polysynthetic language with a high degree of fusion, or to use our new term again, a polyfusional language.

Despite these caveats, the Basque scenario is a good illustration of what is likely to happen when a language is high on both indices. The area of Basque that has been particularly affected by obsolescence is the system of agreement with three arguments, which is increasingly subject to restrictions and gradually being reduced. Basque exhibits both the *Monotransitive* and the *Ditransitive Person-Role Constraint*, albeit to different degrees.

## 5.3 Language contact

Although such a complex agreement system may very well become obsolescent without external interference (or, in particular favourable circumstances, remain intact), the intensive exposure to the pronominal systems of Romance languages certainly accelerated the process. More generally, we observe that the degree of morphological obsolescence correlates with the status of the language. Obsolescence is more advanced in the French part of the Basque Country, where the diglossic situation is still in place, so that practice and transmission of Basque are more difficult. The discrepancy between Lafitte's (1944/2001) grammatical description of Northern Basque from the first half of the 20th century, and the reality of present-day Northern Basque, especially as it is spoken by younger generations, confirms that a large part of the complex verbal morphology has been lost within only a few generations. The higher exposure to formal Basque in the Southern provinces during the past decades has slowed down the process of obsolescence which has been continuing for centuries, whereas the persisting diglossia in the Northern provinces has accelerated this process. As a consequence, Northern Basque is more analytic than Southern Basque (cf. Jendraschek 2003a:40–42). Convergence of Basque with Romance is not the only possible outcome though, as we have seen in §4 some examples of semantic simplification of utterances, i.e. morphosyntactic simplification without compensating strategies. This strategy is independent of language contact and rather differentiates Basque from Romance. In Basque, complex morphosyntactic constructions are indeed either a characteristic of conservative registers or of 'translatese' that calques Romance morphosyntax (Jendraschek 2007:321). Typological evolution goes towards less synthesis in all varieties, whereas standardization (including both corpus and status planning) goes towards conservation of (many, but not all) synthetic forms (Jendraschek 2003a:42). Paradoxically, contact with Romance (and increasingly English) seems to favour both evolutions. On the one hand, there is simplification of argument constellations and agreement patterns in informal speech. On the other hand, language contact contributes to the preservation of discourse patterns typical for Standard Average European, i.e. utterances with marked argument constellations, marked TAM categories, and frequent clause subordination. As such discourse patterns are usually rendered in Basque by

complex morphology, formal registers are predominantly affected. This may explain why most Basques experience a 'register gap' between (informal) spoken and (formal) written Basque. This register gap takes the form of diglossia where dialectal variation has to be taken into consideration (cf. Hualde 2003a:5).

# References

Addis, Kristin. 1993. Paradigm trimming in Basque. *Linguistics* 31.3:431–474.

Aikhenvald, Alexandra Y. 2007. Typological distinctions in word-formation. In *Language Typology and Syntactic Description. Volume 3, Grammatical Categories and the Lexicon.* Second edition, ed. T. Shopen, 1–65. Cambridge etc.: Cambridge University Press.

Altuna, Patxi, and Miren Azkarate. 2001. *Euskal morfologiaren historia.* Donostia: Elkarlanean.

Arregi, Karlos, and Andrew Nevins. 2007. Obliteration vs. Impoverishment in the Basque g-/z- Constraint. In *Proceedings of the 26th Penn Linguistics Colloquium, Penn Working Papers in Linguistics 13.1*, 1–14.

Arregi, Karlos, and Andrew Nevins. 2008. A principle order to postsyntactic operations. URL http://home.uchicago.edu/~karlos/

Bybee, Joan. 2001. Mechanisms of change in grammaticization : the role of frequency. In *The Handbook of Historical Linguistics*, ed. B. Joseph and R. D. Janda 2003, 602–623. Oxford: Blackwell.

Bybee, Joan, and Paul Hopper. 2001. Introduction to frequency and the emergence of linguistic structure. In *Frequency and the Emergence of Linguistic Structure*, ed. J. Bybee and P. Hopper, 1–24. Amsterdam/Philadelphia: Joan Benjamins.

Clyne, Michael. 1992. *Pluricentric languages. Differing norms in different nations.* Berlin & New York: Mouton de Gruyter.

Croft, William. 2000. *Explaining Language Change. An Evolutionary Approach.* London: Longman.

Comrie, Bernard. 1989. *Language universals and linguistic typology.* Oxford & Cambridge, USA: Blackwell.

Dorian, Nancy C. 1978. The fate of morphological complexity in language death: Evidence from East Sutherland Gaelic. *Language* 54.3:590–609.

Dudenredaktion. 2009. *Duden. Die Grammatik.* 8th edition. Mannheim: Dudenverlag.

Echenique Elizondo, Maria Teresa. 1997. *Estudios lingüísticos vasco-románicos.* Madrid: Istmo.

Fernández, Beatriz. 2001. Absolutibo komunztaduradun ergatiboak, absolutibo komunztaduradun datiboak: Ergatiboaren lekualdatzetik datiboaren lekualdatzera. In *On case and agreement*, ed. B. Fernández and P. Albizu, 147–165. Bilbao: EHU.

Haase, Martin. 1992. *Sprachkontakt und Sprachwandel im Baskenland. Die Einflüsse des Gaskognischen und Französischen auf das Baskische.* Hamburg: Buske.

Haspelmath, Martin. 2004. Explaining the ditransitive person-role constraint: A usage-based approach. *Constructions* 2/2004.

Haspelmath, Martin. 2009. An empirical test of the agglutination hypothesis. In *Universals of language today*, ed. S. Scalise, E. Magni and A. Bisetto, 13–29. Dordrecht: Springer.

Haspelmath, Martin & Andrea D. Sims. 2010. *Understanding Morphology.* 2nd edition. London: Hodder Education.

Hualde, José Ignacio. 2000. Bizkaiko kostaldeko hizkerak atzo, gaur eta bihar. In *Dialektologia gaiak*, ed. K. Zuazo, 7–25. Vitoria-Gasteiz: Arabako Foru Aldundia.

Hualde, José Ignacio. 2002. On the loss of Ergative Displacement in Basque and the role of analogy in the development of morphological paradigms. In *The linguist's linguist: A Collection of papers in honour of Alexis Manaster Ramer*, ed. F. Cavoto, 219–230. München: Lincom Europa.

Hualde, José Ignacio. 2003a. Introduction. In *A Grammar of Basque*, ed. J. I. Hualde and J. O. de Urbina, 1–14. Berlin/New York: Mouton de Gruyter.

Hualde, José Ignacio. 2003b. Finite forms. In *A Grammar of Basque*, ed. J. I. Hualde and J. O. de Urbina, 205–246. Berlin/New York: Mouton de Gruyter.

Jendraschek, Gerd. 2003a. When old paradigms die, new paradigms are born: On the eternal cycle of morphological change and its importance for language typology. *Journal of Universal Language* 4.2:27–59.

Jendraschek, Gerd. 2003b. L'évolution diachronique de l'expression de la possibilité en basque. *Lapurdum* 8:263–283.

Jendraschek, Gerd. 2006. Basque in contact with Romance languages. In *Grammars in contact. A cross-linguistic typology.* ed. A. Y. Aikhenvald and R. M. W. Dixon, 143–162. Oxford: Oxford University Press.

Jendraschek, Gerd. 2007. *La notion modale de possibilité en Basque: Morphologie, syntaxe, sémantique, variations diachronique et sociolinguistique.* Paris : L'Harmattan.

King, Alan R. 1993. Communicative grammar of the Basque verb (selected aspects). Doctoral Dissertation, London University.

Lafitte, Pierre. 2001. *Grammaire basque (Navarro-Labourdin littéraire).* San Sebastian/Bayonne: Elkarlanean. Reprint. Original edition, Bayonne: Librairie 'Le livre', 1944.

Lafon, René. 1943. *Le système du verbe basque au XVI siècle.* Donostia/Baiona : Elkar.

Laka, Itziar. 1993. The structure of inflection : A case study in X° syntax. In *Generative Studies in Basque Linguistics*, ed. J. I. Hualde and J. O. de Urbina, 21–70. Amsterdam/Philadelphia: John Benjamins.

Lehmann, Christian. 2002. Thoughts on grammaticalization. Second, revised edition. Universität Erfurt: Arbeitspapiere des Seminars für Sprachwissenschaft der Universität Erfurt Nr. 9.

Matisoff, James A. 1991. Areal and Universal Dimensions of Grammatization in Lahu. In *Approaches to grammaticalization.* Volume II, ed. E.C. Traugott and B. Heine, 383–453. Amsterdam/Philadelphia: Joan Benjamins.

Plank, Frans. 1999. Split morphology: How agglutination and flexion mix. *Linguistic Typology* 3:279–340.

Schrodt, Richard, and Karin Donhauser. 2003. Tempus, Aktionsart/ Aspekt und Modus im Deutschen. In *Sprachgeschichte. Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung.* 2nd edition. Volume 3, ed. W. Besch, A. Betten and O. Reichmann, 2504–2524. Berlin: Walter de Gruyter.

Song, Jae Jung. 2005. Grammaticalization and structural scope increase. *Linguistics* 43.4:795–838.

Trask, R. L. 1997. *The History of Basque.* London/New York: Routledge.

Trask, R. L. 1998. The typological position of Basque: Then and now. *Language Sciences* 20.3:313–324.

Trask, R. L. 2003. The noun phrase: nouns, determiners and modifiers; pronouns and names. In *A Grammar of Basque*, ed. J. I. Hualde and J. O. de Urbina, 113–170. Berlin/New York: Mouton de Gruyter.

Zubiri, Entzi, and Ilari Zubiri. 2000. *Euskal Gramatika Osoa.* Bilbao: Didaktiker.