This article deals with the ways in which non-lexical core arguments can be expressed in various languages. It tries to devise a typological hierarchy for the different types and endeavours to place Romance within this hierarchy. An analysis of Basque verbal markers as cross-reference morphemes introduces the subject with a language radically different from central IE. Using Nichols’ (1986 & 1992) typological differentiation between head-marking and dependent-marking languages as its basis, a typological sub-parameter of “clausal head-marking vs. clausal dependent-marking” is suggested which is shown to correspond to two radically different types of clausal co-reference: (1) agreement (concord) and (2) cross-reference. This terminology is then used to describe and explain an ongoing syntactic change in which Spanish object clitics have evolved into obligatory verbal markers closely resembling those of Basque. Their conventional analysis as “agreement markers” is questioned and Spanish is shown to be moving towards a clausal head-marking language in which all core-arguments of the sentence have to be expressed by verbal affixes, while nominal and pronominal argument realisations become mere appositions outside the sentence core. The traditional concept of an emerging new paradigm of “object conjugation” is rejected.

1. Introduction

Empty categories have traditionally been invoked to account for phenomena like the optionality of surface subjects in pro-drop languages like Latin, Italian, or Spanish and it is contended that e.g. in Italian ti amo there is no phonetic realisation of a referential subject. The descriptive device of empty categories has, however, not been limited to cases of ‘empty’ subject arguments but is also used to explain apparently unrealised object arguments in languages like Brazilian Portuguese (Goldbach 1999) and Basque. The latter has been

1 I’d like to thank Harro Stammerjohann for commenting upon an earlier draft of this paper, Martin Haase for checking my Basque examples and data, Ulrich Detges and Rolf Kailuweit for useful comments on my Saarbrücken talk, and an anonymous reviewer for valuable comments on an earlier draft of this paper. John Cleek (Minneapolis) was kind enough to polish my English style and punctuation. Many improvements on this text can be credited to them, any shortcomings, of course, are exclusively due to my own intellectual limitations.
considered to be a ‘null object language’ on the basis that objects need not be realised on the surface, neither by lexical NPs nor by Pronouns. Goldbach writes:

Languages like Pashto and Basque, which exhibit agreement with the subject as well as with the object but which do not license null topics, possess null subjects and null objects (translated from Goldbach 1999:64).

While this empty category hypothesis has been widely discussed with respect to the more thoroughly researched IE languages of Europe, I feel, however, that especially the Basque case calls for further discussion. In the second half of this paper, the results of this discussion will then be used to shed light also on the much more well studied languages like e.g. Spanish. Applied to Basque, the null subject and null object hypothesis has always struck me as particularly counterintuitive because it doesn’t seem to account correctly for the facts of this language’s verbal morphology. Basque possesses a complex machinery of verbal inflection that serves only one purpose: to assure the overt surface realisation of every single argument in a sentence. Actually, sentences with effectively unrealised objects are exceedingly rare in Basque and seem to occur only in connection with nominalisations like the following, where there simply are no morphological slots on the nominalised element to which non-lexical objects could possibly be attached:

(1) Zakur-rengana                  abiatu  da
    dog-ALATIVE.ANIM.DET.PL       proceed  aux.1.SG.PRES.ABS
    lot-ze-ko.\textsuperscript{2}
    tie_up-NOMINALISER-in_order_to
    “He went to the dogs to tie (them) up.”

Basque might therefore be described as having a radical dislike for any kind of non-realisation of object-arguments. However, the general agreement among generative scholars turns out to be diametrically opposed to this analysis and makes Basque a language with far-reaching object-drop capabilities. The question whether Basque either drops its objects regularly or else is reluctant to ever drop them at all is the starting-point for the following considerations: How is it possible to see the same state of affairs so differently?

It may prove insightful to set this question in the larger context of the strategies to be found in various languages for the realisation of non-lexical subject and object arguments, i.e. I shall take an onomasiological stance on a universal syntactic problem. In particular, I shall consider the question of drawing a boundary between the concepts of ‘agreeing with’ and ‘instantiating’ an argument in various language types. There appears to be a continuum of

\textsuperscript{2} This is a sentence I found in a teach-yourself Basque manual (Beaumont & Lazkano 1998:111).
referential forms used by different languages for the expression of non-lexical arguments which ranges from morphologically free forms like German or English personal pronouns, via clitics (most Romance languages) to the polysynthetic realisations of cross-reference languages like Basque. I shall argue that the features to be found in certain modern Romance languages (essentially Spanish, Catalan, Galician, and European Portuguese) were best explained as a typological shift towards a type quite similar to that of Basque. As this is a generally much less familiar language than the others to be discussed later, I feel that it is important to dwell rather more extensively on a description of those basic facts of Basque syntax I will later cite to analyse the ongoing typological change within Romance.

2. The Realisation of non-lexical arguments in Basque

Basque is a language isolate showing ergative morphology, i.e. subjects of intransitive (or rather: unaccusative) verbs as well as direct objects are marked for absolutive case, whereas subjects of transitive verbs take ergative case marking. Basque possesses only very few verbs capable of taking synthetic inflection-affixes: according to the authoritative Basque grammar of the Basque Language Academy \textit{Euskaltzaindia} from 1987, out of originally over 60 synthetical verbs in the earliest surviving texts, only 26 are still in modern usage and even fewer if we exclude those which are limited to the literary language (cf. Gómez & Sainz (1995:238f.). Most verbal concepts can only be expressed periphrastically with the aid of an auxiliary.

The small number of synthetic verbs in combination with a morphology of the agglutinating type, i.e. a verbal morphology with relatively little allomorphy and amalgamations, allows Basque verbs to co-refer not only with their subjects, but also with all other arguments, direct and indirect objects alike. There is even a fourth type of co-reference affix, called ‘allocutive’ referring to second person non-arguments of the ‘ethical dative’-type, which I shall make no further mention of as we will only be concerned with the realisation of core-arguments (cf. Gómez & Sainz 1995:236). Moreover, Basque has consistently been classified as a pro-drop language in generative accounts, given that nominal and pronominal subjects appear to be freely elidable. How then do all these elements interact in the construction of intransitive, transitive, and ditransitive sentences?
2.1 Basque, one argument (intransitive, Nor)

By way of exemplification, let us first analyse the intransitive Basque sentence *Ni abiatzen naiz* “I set out”, *Zu abiatzen zara* “You set out” etc. as follows:

| 1 sg.   | (ni-ø)_i | naiz_i |
| 2 sg.   | (zu-ø)_i | zara_i |
| 3 sg.   | (hura-ø)_i | da_i |
| 1 pl.   | (gu-ø)_i | gara_i |
| 2 pl.   | (zuek-ø)_i | zarete_i |
| 3 pl.   | (haiek-ø)_i | dira_i |

| pron-ABS | set_out-PTCP.PRS | aux be.PRS.IND.ABS1SG … |
| “[I / you / he she / we / you / they] set(s) out.” |

The example shows absolutive case marking on the facultative subject pronoun (*ni, zu, hura* etc.), a non-finite verb-form carrying lexical and aspectual information, and the auxiliary *izan* “to be”, marked for tense and mood and co-referenced with the subject pronoun. Utterances like *abiatzen da* “he sets out” and *abiatzen gara* “we set out” without a subject noun phrase or pronoun nevertheless constitute full sentences. Like the pronouns, nominal 3rd person subjects as well carry absolutive case which is marked by a ø-ending:

(2)  
*Jon-ø abiatzen da.*  
*Jon-ABS parting is*  
*“Jon is setting out.”*

2.2. Basque, two arguments S-DO (transitive, Nor-Nork)

In Basque transitive utterances with two arguments of the type ‘He sees the book’, absolutive case marking goes to the direct object whereas the transitive subject receives ergative case. In this way, subjects of transitive sentences are morphologically distinct from those of intransitive sentences. Again, the verbal concept (here ‘to see’) is realised periphrastically; in transitive sentences, the

---

3 This is the traditional name for the construction in Basque grammar. *Nor, nori* and *nork* are the case-marked forms of the interrogative “who” which are used as the Basque names for absolutive, dative and ergative as well as for the sentence constructions in which they participate.

4 The zero-endings marked as ø in the morphological analysis represent slots in the paradigm in which the absence of an affix is one of a set of possible values within the paradigm; ø therefore contrasts with the presence of other affixes and marks a zero-ending, not the zero-realisation of an argument (= Ø). These morphological facts have nothing to do with the descriptive device in syntactic analysis of positing empty categories like traces or *pro* and the like.
auxiliary has to be *ukan* ‘have’, rather than *izan* ‘be’, as only *ukan* has the necessary cross-reference capabilities which allow co-reference not only with the ergative-subject but also with the absolutive-direct object.

<table>
<thead>
<tr>
<th>Case</th>
<th>Affix</th>
<th>Inflected Verb</th>
<th>Inflected Verb Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 sg.</td>
<td>(ni-k)</td>
<td><em>ikan</em></td>
<td><em>ikan-ten ditu-ti</em></td>
</tr>
<tr>
<td>2 sg.</td>
<td>(zu-k)</td>
<td><em>liburu-a</em></td>
<td><em>liburu-ak ditu-ziu</em></td>
</tr>
<tr>
<td>3 sg.</td>
<td>(har-k)</td>
<td><em>ikan</em></td>
<td><em>ikan-ten ditu-ø</em></td>
</tr>
<tr>
<td>1 pl.</td>
<td>(gu-k)</td>
<td><em>ikan</em></td>
<td><em>ikan-ten ditu-gi</em></td>
</tr>
<tr>
<td>2 pl.</td>
<td>(zue-k)</td>
<td><em>ikan</em></td>
<td><em>ikan-ten ditu-zuei</em></td>
</tr>
<tr>
<td>3 pl.</td>
<td>(haie-k)</td>
<td><em>ikan</em></td>
<td><em>ikan-ten ditu-ztei</em></td>
</tr>
</tbody>
</table>

Figure 2: Basque, two arguments (S + DO-3 Sg)

Somewhat simplifying the morphological situation, one might analyse the affix *du-* as co-referencing the singular direct object while *ditu-* corresponds to a direct object in the plural. At the same time, *du-* / *ditu-* are also marked for 3rd person, as there are special forms for 1st and 2nd person as well:

(3) (Ni-k) *ikan* *zai-tu-ti.*
I-ERG.SG see-PTCP.PRS ABS.2SG-aux have.PRS.IND.ERG.1SG
“I see you (sg.).”

(4) 

You (sg.) see me”, etc.

Again, the inflected verb group on its own already constitutes a complete sentence, as all lexical arguments are freely elidable:

(4) *ikan* *di-tu-ti.*
see-PTCP.PRS ABS.3PL-aux have.PRS.IND.ERG.1SG
“I see them”.

Similar affixes – and similar one-word sentences – also occur with the other synthetic verbs like e.g. *jakin* ‘to know’:

(5) a. *jakin* *d-aki-t.*
reason-DET.ABS.SG ABS.3PL-know.PRS.IND.ERG.1SG
“I know the reason.”
Ba-d-aki-t.  
ENUNC-ABS.3PL-know.PRS.IND-ERG.1SG  
“I know it.”

b. Arrazoi-ak d-aki-zki-t.  
reason-DET.ABS.PL ABS.3PL-know.PRS.IND-ABS.3PL-ERG.1SG  
“I know the reasons.”

Ba-d-aki-zki-t.  
ENUNC-ABS.3PL-know.PRS.IND-ABS.3PL-ERG.1SG  
“I know them.”

c. Arrazoi-a d-aki-gu.  
reason-DET.ABS.SG ABS.3SG-know.PRS.IND-ERG.1PL  
“We know the reason.”

Ba-d-aki-gu.  
ENUNC-ABS.3SG-know.PRS.IND-ERG.1PL  
“We know it.”

d. Arrazoi-ak d-aki-zki-gu.  
reason-DET.ABS.PL ABS.3PL-know.PRS.IND-ABS.3PL-ERG.1PL  
“We know the reasons.”

Ba-d-aki-zki-gu.  
ENUNC-ABS.3PL-know.PRS.IND-ABS.3PL-ERG.1PL  
“We know them”.

It is this type of structures that has given rise to the description of Basque as a null object language. Actually then, Basque is not only considered to be pro-drop but has also been dubbed an ‘object-drop-language’ (Landa & Franco 1996:160). Saltarelli goes even one step further and characterises Basque as a ‘null argument language’ in which any type of lexical argument is always optional as far as syntax is concerned:

[...] with respect to the issue of empty categories, Basque is a null subject language, as Italian has been defined. Moreover, Basque is a null object language (direct and indirect). If we follow Taraldsen’s original hypothesis that null subject languages are a consequence of the richness of the verb inflectional properties of the language, then Basque is properly a ‘null argument language’ since its verb inflectional parameter marks subject (both thematic and non-thematic), direct object and indirect object (Saltarelli 1988:XIX).

2.3. Basque, three arguments S-IO-DO (transitive, Nor-Nori-Nork)

By adding an indirect object as a further argument, we get the ditransitive construction of the type “He gives the man the book” which is rendered in Basque by marking the subject as ergative, the direct object as absolutive, and the indirect object as dative. Again, all arguments are cross-referentially represented on the auxiliary ukan:
As can be gathered from the gloss, the auxiliary now has to carry an
absolutive marker referring to the direct object, a dative marker referring to the
indirect object and an ergative marker, securing co-reference with the subject.
Again, the table only gives a glimpse of the actual possibilities and intricacies
of Basque cross-reference morphology. A sentence like (6) below:

(6) \((nik)\_gizon-ari\_liburu-a\_ema-ten\)

shows a reflex of plurality on the direct object and on the auxiliary, if the direct
object is in the plural:

(7) \((nik)\_gizon-ari\_liburu-ak\_ema-ten\)

Changing the indirect object to plural as well, we get (8):

(8) \((nik)\_gizon-ei\_liburu-ak\_ema-ten\)

As can be seen in these examples, Basque is markedly different from the
central Indo-European languages in that its finite verbs obligatorily reflect in
their morphology all basic relationships of argumenthood in a sentence, a
syntactic property known as *cross-reference* in language typology. The status of
these object affixes has been described in various ways in the literature:
Schwerteck (1984:9), in talking about the incorporating verb forms, calls the object affixes “pronouns”, “segmental markers” und “pronominal elements”. This terminology reflects the intuition that these affixes themselves carry out syntactic functions which go beyond mere agreement; calling them pronouns, however, would endow them with a degree of morphological independence which they definitely do not possess.

Franco (1991 and 2000) talks about “verbal agreement morphemes”, Goldbach (1999:10) about a “system of object agreement” and “inflectional endings on the verb”. This is standard generative terminology which takes into account their stronger morphological integration into the verb, but sees them as mere agreement devices.

Saltarelli (1988:XVII) generally uses the unspecific term “marker”; he does, however, call Basque a “highly inflected language” and seems to place less weight on the agglutinative character of Basque morphology.

Most authors seem to concur then in considering these markers ‘agreement affixes’ of some kind and the following analysis by Landa and Franco represents a generally accepted view concerning the structure of Basque sentences. As in English or Spanish, verbal markers are seen as only co-indexing the real (lexical or pronominal) arguments. If these happen to be absent, this phenomenon is seen as ‘elision’ and the index goes to an empty category – a trace or a pro:

(9) Jonek aulkia apurtu du eta nik konpondu dat.

Jon-ERG chair-ABS break ABS.3SG-AUX have-ERG.3SG and I-ERG
konpondu d-u-t
fix ABS.3SG-AUX have-ERG.1SG
“Jon has broken the chair and I have fixed (it).” (Landa & Franco 1996:161)

---

5 “Pronomina [...] segmentale Bedeutungsträger [...] pronominale Elemente” in the German original. As to the historical origin of these polysynthetic verbal complexes of Basque he writes: “It is by no means new that Basque verbal forms, when retraced to former states in their history, have been considered to be entire sentences ([Footnote 21:] Examples for this view can be found e.g. in H. Schuchardt’s *Primitiae linguae vasconum*)” (translated from Schwerteck 1984:15).

6 “Just as the null-subject possibilities of Italian and Spanish follow from the potential of their subject inflection, the null-object possibilities of Pashto and Basque result from their system of object agreement” (translated from Goldbach 1999:65f.). This is, of course, an example for what Haspelmath (1999:188) has dubbed the “teleological fallacy”: rich agreement morphology may be an important factor in facilitating the described phenomena but can not in itself be seen as causing it.

7 “[...] like for example in Basque, where inflectional endings unequivocally indicate the possible subjects (or objects) of a verb” (translated from Goldbach 1999:10).
According to this theory, the direct object of the coordinated sentence can be elided easily because the object agreement markers on the verb identify the null object completely. They co-index the direct object *aulkia* “chair” with the empty constituent in the object position of the coordinated sentence. Note, however, that this is by no means the only possible analysis; after all, both auxiliaries already contain a marker for 3sg. absolutive and we might envisage an alternative view in which the observable verbal affixes, rather than a hypothetical Ø, serve as the locus for co-indexation with the lexical direct object as in (10):

(10) Jonek *aulkia* i apurtu d_u-ø
Jon-ERG chair-ABS break ABS.3SG-AUX have-ERG.3SG
eta nik konpondu d_u-t.
and I-ERG fix ABS.3SG-AUX have-ERG.1SG
“Jon has broken the chair and I have fixed (it)”

Why use an empty constituent if there are actual surface elements available that could be seen either as agreement markers or just as well as pronoun-like representations of the direct object itself? As a consequence of the non-obligatory nature of Basque NP-arguments on the one hand and Du Bois’ *Preferred Argument Structure constraint* (“Avoid more than one lexical core argument” cf. Du Bois 2003:34ff.) on the other, the great majority of all arguments would then have to be treated as not occurring on the surface, triggering an immense proliferation of empty syntactic categories. This is likely to be an unnecessary violation of any principle of descriptive economy and therefore a classical candidate for an application of Occam’s razor.

A further relevant observation about this “empty constituent” can be gained from a comparison with null objects in other languages like e.g. Spanish. A major difference between Basque and Spanish lies in the fact that in Basque the only possible alternative to the null object would be either the full lexical noun phrase *aulkia* “chair” itself or else an emphatic demonstrative. This is due to the (highly noticeable!) fact that Basque, unlike Spanish, does not possess a class of non-demonstrative object pronouns. Nevertheless, the direct object does not have to be contextually inferred; rather, it is unequivocally specified in the verbal morphology as to case and number, namely absolutive singular. That is, the object affixes are specified with the same or even more precision than the corresponding Romance clitics or the German pronouns. One might ask, therefore, why Basque verbal affixes are usually not considered to be realisations of arguments but are seen rather as mere agreement morphology referring to lexically realised arguments or their respective traces.

These elements are clearly different from such well-established agreement markers as e.g. subject agreement morphemes in German or English. In these languages, verbal markers do not only agree with noun phrase subjects but, in
their absence, also with a class of pronominal subjects which in non-pro drop languages as German are even obligatory:


This is prototypical ‘agreement’: morphological co-reference with an overt NP or pronoun. Basque, however, is different. While first and second person indirect objects might still be overtly realised as case-inflected forms of emphatic personal pronouns, no such option is available for third person objects. Basque only has 1st and 2nd person genuine personal pronouns; it is only for the sake of completeness that grammars tend to list the demonstratives *hura* and *haiek* for the function of emphatic 3rd person pronouns (cf. Zubiri 2000:47). Therefore, if we consider the verbal markers to be mere agreement morphemes, the least emphatic way of expressing third person pronominal objects on the surface would be demonstratives. This syntactic fact casts further doubt on the analysis of Basque verbal affixes as agreement markers, because if these elements actually were agreement affixes, the question arises: What are they are supposed to be agreeing with?

Let’s finally turn to a last example in which the auxiliary is not only marked for cross-reference but also for tense and mood, and which is therefore better suited to show the full possible complexity of Basque finite verbs and the agglutinating character of the language (taken from Saltarelli (1988:XVII)):

(12)  *Guk gizonei liburuak eman diezazkiegu.*

```
Guk     gizon-ei      liburu-ak       ema-n
we-ERG man-DAT.DAT.PL book-DAT.ABS.PL give-PRTC.PERF

d       -i           -eza         -zki     -e     -ke       -gu
ABS.3PL-DAT.3PL -AUX have.PRS -ABS.3PL -DAT.3PL -MOOD -ERG.1PL
```

“We can give the books to the men.”

All lexical arguments have to be co-referenced by a marker on the auxiliary:

– The subject, i.e. the ergative pronoun *<guk>* by the marker ERG.1PL *<-gu>*.
– The dative indirect object *<gizonei>* by the discontinuous dative marker *<-i->* for 3rd person and *<-e->* for plural.
– And finally, the absolutive direct object *<liburuak>* by the discontinuous absolutive marker *<d->* for 3rd person and *<-zki->* for plural.

The complete representation of all arguments on the auxiliary allows for any lexical ergative, absolutive, and dative argument to be optional and the verbal complex alone constitutes a complete grammatical utterance:
Eman diezakiekengu.
Ema-n     d            -i             -eza                -zki          -e
give-PTC.PERF  ABS.3PL-DAT.3PL  -AUX.PRS  -ABS.3PL  -DAT.3PL
-ke        -gu
-MOOD  -ERG.1PL
“We can give them to them.”

The situation in Basque is thus radically different from that in German or, if slightly (and sometimes even considerably) less, from that of the Romance languages. Now, there may be the danger of illicitly projecting our intuitions gained from more familiar Indo-European languages onto the situation in a non-Indo-European language like Basque, for which reason the following two questions are by no means trivial:

– Are the ergative, dative and absolutive noun phrases, guk, gizonei and liburuak, really the subject, indirect, and direct object of the sentence, in the same sense in which the nominative, dative, and accusative arguments in the German sentence Wir können den Männern die Bücher geben (“We can give the books to the men”) would be said to instantiate the subject, indirect object, and direct object?

– Are the Basque cross-reference affixes really only agreement markers, comparable to verbal agreement in German Wir können (“We can”)?

An answer to these questions is not only important to Bascologists but could also aid us in reassessing the well-known facts of the Romance languages, shedding light on the syntactic status of clitics in those constructions that have sometimes been called “clitic-doubling” and sometimes “Romance object conjugation”.

3. The typological parameter head-marking vs. dependent marking

In order to do justice to languages like Basque, Johanna Nichols in her 1986 seminal paper on head-marking vs. dependent-marking languages introduced a major typological distinction that may well be relevant to the problem of non-lexical argument realisation in various languages. Nichols draws our attention to an important empirical imbalance in linguistic research that existed at that time (and most probably persists until today): while head-marking features are statistically very important and probably even majoritary among the languages of the world (Nichols 1986:89ff.), this language type had received only very little if any attention by linguists – including those who claimed to be working on language universals:

8 In this paper she discusses a sample of over 60 typologically diverse languages, among which there is – quite surprisingly – not one Romance language! The author developed her theory further in Nichols (1992).
Despite the efforts of formal grammarians to take a range of languages into consideration, there is a glaring gap in the typological coverage: the exotic languages that have so far received significant attention have been almost exclusively dependent-marking (Japanese, Korean, Finnish, Malayalam, Australian languages) or double-marking (Turkish, Arabic, Hebrew, Australian languages). Of the head-marking languages, only Navajo has received significant theoretical attention. In addition, Relational Grammar has investigated a number of head-marking languages (primarily from the Algonkian, Salishan, and Wakashan groups); but these languages have not been used to raise questions of constituency, centricity, syntactic bonds, government etc. (Nichols 1986:115f.).

Nichols’ typological analysis is based on two universal principles, namely headedness and overt morphological marking. Headedness is meant to convey the notion that natural language constituents regularly come as structured into a head, an obligatory constituent determining the overall syntactic status of the whole constituent, and one or more facultative dependents which derive their morphosyntactic accidents from the head. This type of configuration can be shown to exist not only at the phrase but also at the clause and sentence level:

<table>
<thead>
<tr>
<th>Constituent</th>
<th>Head</th>
<th>Dependent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun phrase</td>
<td>possessed noun</td>
<td>possessor</td>
</tr>
<tr>
<td></td>
<td>noun</td>
<td>modifying adjective</td>
</tr>
<tr>
<td>adposition</td>
<td>object of adposition</td>
<td></td>
</tr>
<tr>
<td>Clause</td>
<td>predicate / verb</td>
<td>arguments / adjuncts</td>
</tr>
<tr>
<td>auxiliary verb</td>
<td>lexical / main verb</td>
<td></td>
</tr>
<tr>
<td>Sentence</td>
<td>Main-clause predicate</td>
<td>subordinate clause</td>
</tr>
</tbody>
</table>

Figure 4: Head-dependent-configurations (cf. Helmbrecht 2001:1425)
If the phenomenon can occur independently on different levels, it should be legitimate to extract typological sub-parameters from the general theory. I therefore suggest a typological parameter *clausal head-marking* to be derived and abstracted from the general principle that Nichols advocates for. As to the distribution of head and dependent on the clause level, Nichols argues (1986:57) that predicates constitute the head, whereas arguments and adjuncts can be analysed as dependents.9 Clausal head-marking languages, then, are those in which all arguments of the sentence are usually realised by morphological means on their finite verbs or auxiliaries. I contend that the question of non-lexical argument realisation can be seen along the lines of clausal head-marking vs. clausal dependent-marking and that Nichols’ parameter may be used to shed new light on the historical development of clitic systems in Romance in comparison with languages like Basque or German which represent markedly different types.

Due to the fact that the tendency for consistently marking either on the head or on the dependent is not absolute, there are actually further types to be found: in *double-marking*, syntactic relations are redundantly marked on heads and dependents alike, whereas in *split-marking* languages, the situation may be different, depending on the grammatical subsystem. Basque is actually a good example for both types: at the sentence level, all syntactic relations are reflected on the verb (head) but argument noun phrases are also marked for case so that we get clausal double-marking. Moreover, Basque belongs to the split-type, as double-marking is limited to the clause level in this language, which is otherwise largely dependent-marking on the phrasal level.

For an example of a genuine *head-marking* language we may want to take a look at Abkhaz. Argument noun phrases in Abkhaz bear no case markers whatsoever and the entire morphosyntactic information is polysynthetically realised on the verb, which morphologically cross-references with the lexical arguments for person, number, and gender. In the absence of morphological case, the syntactic relations are encoded through the serialisation of verbal

---

9 In other cases, auxiliaries and their respective lexical (‘main’) verbs may also stand in a clausal head-dependent relation to each other. I will not dwell further on this interpretation.
affixes. All argument noun phrases are optional and appear only if they are discourse-pragmatically needed to establish the reference of the verbal affixes.

(14) Pure head-marking language: Abkhaz

\[ a-xaç'a \quad a-ph'os \quad a-sëqô 'q \quad ø-l-y-te-yt' \]

the-man the-woman the-book it-to_her-he-gave-FINIT

“The man gave the woman the book.” (Helmbrecht 2001:1427)

The inflected verb group again constitutes a complete sentence, quite similar to the situation in Basque, as we have already seen:

(15) \[ ø-l-y-te-yt' \]

it_to_her-he-gave-FINIT

“He gave it to her.”

Now, Nichols observes that this effect seems to be characteristic of all languages with head-marking morphology at the clause level. This is hardly surprising, as these languages code the entire grammatical information about the clausal arguments like case, number, gender etc. on the verb, leaving noun phrases with the sole function of providing lexical information about new discourse participants. Speakers are free to omit lexical or pronominal material if it is referentially recoverable in the discourse universe. In these languages, the surface appearance of lexical noun phrases and pronouns is bound exclusively to discourse motivations and can therefore be completely optional at the syntactic level. When full noun phrases appear at all, they behave like appositions with respect to their coreferential verbal affixes and are realised outside the sentence core.

In head-marked grammatical relations, the dependent is usually an optional element of the constituent. For instance, in languages with consistently head-marked clauses, the verb itself normally constitutes a complete sentence; full NP’s are included only for emphasis, focus, disambiguation etc. (Nichols 1986:107).

So, in head-marking languages there is a sharing of functions between the verbal affixes and the full noun phrases which is quite comparable to that between full verb and auxiliary in verbal periphrases, in that one element, the auxiliary, specialises in conveying only grammatical information while the lexical information is left entirely to the other element, the full verb. In the same vein, head-marking verbal affixes specialise in the purely grammatical

10 “The appositional status of NPs in Abkhaz can be found in many other non-European languages and may be viewed as a defining feature, among others, of polysynthetic languages” (Helmbrecht 2001:1427).

11 That lexical noun phrases should be analysed as appositions (or ‘adjuncts’ or ‘dislocated elements’ respectively) in these cases has also been suggested by generativist authors. Aoun (1993) proposes such an interpretation for Lebanese Arabic (cf. the critical discussion in Franco 2000:152ff.).
function while the corresponding appositions carry the lexical information, may receive contrastive stress and can be freely positioned in the sentence periphery according to the requirements of discourse pragmatics. The parallelism is, however, not complete, as in verbal periphrases both elements are obligatory while in head-marking languages only the verb complex is. In Basque and Abkhaz sentences, the verbal affixes already seem to constitute the arguments while the full noun phrases show all the same behaviour of facultative appositions.

Nichols was by no means the first to propose this analysis; rather it seems to have been the view of almost any linguist ever to have worked on American Indian languages:

The dependent in a head-marked constituent stands in a roughly appositive relation to the head (or, more precisely, to the coreferential marker on the head); the term ‘government’, developed by traditional grammar on the basis of exclusively dependent-marked relations [...], is not appropriate for head-marked constituents. Since the appearance of Boas 1911, descriptions of American Indian languages have insisted that subject and object in these languages are in apposition to the pronominal markers on the verb, rather than (as in Indo-European) being syntactically governed by a verb which agrees with them (Nichols 1986:107).

One of the most prominent advocates for this typological distinction is co-founder of Role and Reference Grammar (RRG) Robert Van Valin, who was drawn to it by his work on the head-marking language Lakhota. This may be the reason why RRG is one of the few grammatical theories on whose very foundations a systematic distinction between head-marking and dependent-marking was implemented. In a short introductory manifesto on RRG, Van Valin writes:

In head-marking languages like Lakhota, the bound pronominals on the verb are considered to be the core arguments; overt NPs are within the clause in apposition to them (Van Valin 2005:3)

According to this analysis, Basque verbal affixes would no longer have to be considered agreement markers but rather ‘core arguments’ in themselves; overt noun phrases, on the other hand, would regularly be seen as appositions, which would not only account for their unlimited elidability but also for the lack of syntactic constraints on their position relative to the core-sentence and to each other.

The fundamental typological difference between these two marking-types should also be reflected in the terminology that is used to describe the different type of clausal co-reference that can be observed in head-marking and dependent-marking languages. Most authors, however, tend to carry over the notion of ‘agreement’ also to head-marking languages as well, thus obfuscating the fundamental differences between the two types. This is unnecessary as the
terminological clarification has been around for over seventy years and, what is more, was published in Bloomfield’s easily accessible, classical monograph “Language”. Here Bloomfield (1933:191ff.) introduces the term cross-reference to refer to the type of clausal co-reference which is typically found in head-marking languages:

<table>
<thead>
<tr>
<th>Agreement (= concord)</th>
<th>Cross-Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>dependent-marking</td>
<td>head-marking</td>
</tr>
<tr>
<td>the argument (= dependent) is realised as a full NP or as a pronoun and is morphologically marked (dependent-marking)</td>
<td>there is co-reference with a non-obligatory full lexical NP which syntactically behaves like a non-core constituent (apposition)</td>
</tr>
<tr>
<td>the verb (head) carries a redundant agreement marker</td>
<td>the argument is realised by an obligatory affix on the verb (head-marking)</td>
</tr>
<tr>
<td>example: subject agreement in German or English</td>
<td>example: Basque, Abkhaz; tendencies in Romance, e.g. French Jean où est-il? (Bloomfield 1933:193)</td>
</tr>
</tbody>
</table>

Figure 6: Two types of clausal co-reference

Thus, whereas clausal co-reference is morphologically realised as ‘agreement’ in languages with clausal dependent-marking, the corresponding morphological process in clausal head-marking languages is ‘cross-reference’ – and therefore radically different. If it is true that in language typology it is not important what can be but rather what must be expressed in a language, then Basque definitely has to be described as a language in which all objects have to be expressed on the surface. And it is as far as can be from a null-subject, null-object or null-argument language. Rather, null-arguments are impossible in Basque because the verbal morphology requires the specification of all arguments. Lexical noun phrases, which by analogy with dependent-marking languages could be mistaken for subjects and objects, do not belong to the sentence-core and must be viewed as non-obligatory appositions providing additional lexical and referential information about the arguments they are co-referenced with. Rather than being arguments they only partake in the argumenthood of the verbal affixes.

It should be noted that the typological importance of the parameter ‘clausal head-marking’ derives largely from the cross-reference morphology on the verb that comes with it. Whether lexical and pronominal arguments carry additional case marking, as e.g. in Basque, or at least ‘differential object marking’ (cf. Bossong 1991), as e.g. in Spanish, is without consequence in this context. On the clause level we can therefore class head- and double-marking languages together and oppose them with the dependent-marking languages and their agreement morphology as the only genuinely different type.
If with cross-reference the traditional terminology of “agreement” vs. “argument” is misleading and inappropriate, which terms should be used then in describing the phenomena adduced above? Du Bois (1987), working on the ergative cross-reference Mayan language Sacapultec, provides us with a terminology which most adequately captures the syntactic facts in languages like Sacapultec and Basque. Where dependent-marking languages tend to have two separate things, i.e. arguments and agreement, head-marking languages only have ‘mentions’ which Du Bois defines as a reference item complex consisting of either a bound form alone (a cross-referencing affix [...]), or an overt free form (full NP or independent pronoun) plus its cross-referencing bound form within the same clause (Du Bois 1987:813).

A ‘mention’ then can consist of an obligatory cross-reference marker alone, in case the lexical noun phrase remains un-expressed. If, however, a lexical NP is present, the NP plus the verbal marker together instantiate one discontinuous argument realisation, bound together by cross-reference. This notion acknowledges the fact that in these languages the grammatical and the lexical aspect of argumenthood are typically split up, only the grammatical element being obligatory. We will have to see whether the notion of ‘mention’ might not be better suited to account for certain Romance languages notorious for the problem of determining the status of clitics and affixes.

In this conception, the verbal affixes in Basque constitute actual instances of the arguments with an independent external reference. I have to thank an anonymous reviewer for pointing out to me a number of conventional tests for argumenthood which, apparently, would all have to yield an affirmative result in order to substantiate this hypothesis. These tests include the questions:

i. Can these affixes be the answer to a question?
ii. Can these affixes carry the main accent?
iii. Do these affixes allow for ellipsis of the inflected auxiliary?
iv. Do these affixes combine with other DPs or NPs?
v. Are these affixes flexible in their position with respect to each other and the verb?

Quite clearly, the answer to all the above questions is no. This does, however, not necessarily invalidate these elements’ claim to argumenthood. All the above tests are diagnostics designed to detect arguments in dependent-marking languages in which arguments must be realised as free morphemes. But, as detailed above, it is precisely the nature of cross-reference languages that they realise arguments not as free morphemes but rather as ‘mentions’ which, by definition, can consist of just a verbal affix. In their mentions we find a repartition of function between argument-realising verbal affixes and their facultative appositional counterparts; the affixes fail all the above tests because the tests rely on those syntactic circumstances which would call for the
otherwise elided appositions. The failure of Basque verbal affixes to comply with the traditional tests for argumenthood can therefore not be adduced to disprove their status as arguments in a head-marking language.

4. The other end of the hierarchy: German

At the other end of the hierarchy of clausal head-marking we find languages like German. Its only head-marking feature on the clausal level is concord, i.e. the typical Indo-European agreement of verbs with their subjects. Actually, according to Nichols, pure clausal head-marking seems to be the exception rather than a statistically relevant type, and we may therefore disregard subject agreement and treat German for all practical purposes as a dependent-marking language. In German, only full noun phrases and pronouns can function as arguments:

(16) *Ich gebe Günther das Buch.*
    I.NOM give Günther the.ACC book.
    “I give Günther a book.”

    I give it him / I give him that /
    “I give it to him.”

How free or how bound these pronouns may be in each case, i.e. how far cliticisation has already progressed in German, is not relevant in our context. It is, however, safe to say that German pronouns are syntactically much closer to lexical noun phrases than any ‘weak’ or even ‘strong’ pronoun in the Romance languages (with the possible exception of Brazilian Portuguese, which seems to be moving away from the rest of Romance in this respect).

In the context of the head-marking morphology in Basque, I asked whether we wanted to analyse the verbal affixes as actual realisations of or as mere agreement markers with the subject, direct and indirect object. Irrespective of the answer to this question, no one can deny that such a question might at least be raised for Basque. For German, however, it would make no sense at all:

– The subject markers on German verbs are clearly agreement affixes referring to an independently realised, overt obligatory subject – a full noun phrase or a subject pronoun.
– Object clitics, pronouns, and noun phrases clearly constitute the realisations of the respective arguments and are typically not appositional.

---

12 In this respect, English would be an even better example, as concord has been reduced to a rather precarious agreement in 3rd person singular verbs. But if English has almost no head-marking, it also has almost no morphological (case-) marking on the dependents, which makes it a completely different type.
phrases and pro-forms are in complementary distribution and there can be no *clitic doubling*.

(18) a. *Ich gebe es ihm.*
    I give it him
    “I give it to him.”

b. *Ich gebe es ihm das.*
    I give it him that

c. *Ich gebe es ihm Peter.*
    I give it him Peter

d. *[Mir / *Ich] mir gefällt das.*
    [to_me / I] to_me appeals that
    “This appeals to me / I like this.”

5. The Romance languages – problematic intermediate cases

We have now examined one example of a language with full clausal head-marking and another one with full dependent-marking morphology. The best way to characterise the Romance languages would be to locate them on an abstract scale on different points in between these two more extreme examples. This intermediate status of Romance has already been noticed by Nichols, who explicitly singles out the Romance case as an example of ongoing typological change within the group of the otherwise quite consistently dependent-marking Indo-European languages:

Indo-European has retained its basic type – dependent-marked with subject inflection on verbs [...] for languages preserving the inherited morphology – for some 6,000 years, with only a recent trend toward head-marked clauses in the pronominal clisis of the Romance languages (a process which occurs only after most of the morphology has been lost) (Nichols 1986:89).

This change is driven by a general Romance tendency to introduce new, alternative head-marking strategies. Historically, most pronouns were reduced to pronominal clitics and these have frequently been further amalgamated into affixes and affix groups which increasingly resemble the cross-reference morphology of Abkhaz, Lakhota or Basque, especially in those Romance languages in which obligatory clitic-doubling is gaining ground.

Like most other Indo-European languages, the Romance languages all possess subject co-reference morphology on their verbs, which has traditionally been considered to be agreement just as in the case of languages like German. This view, however, is not undisputed and thus Corbett (1994) finds it disturbing to speak of agreement in a situation where the non-realisation of the purported referent of the agreement is the statistically normal, unmarked case:
Another common problem, and a serious one in some theoretical frameworks, is the existence of so-called ‘pro-drop’ languages. It may be desirable to say of certain languages, for example, Serbo-Croatian, that verbs agree with their subject in person and number, as in ja čitam ‘I read’, ty čitaš ‘you read’. In most circumstances, the more natural utterances would be čitam and čitaš. The problem is that the pronominal agreement controller is normally not present (Corbett 1994:57).

Kailuweit’s (2005) position within the framework of Role And Reference Grammar is more differentiated, as he considers only part of the cases as instances of agreement. Kailuweit no longer analyses the subject markers on Spanish verbs as agreement affixes as has been traditionally done. On his account, at least as long as no lexical or pronominal subject is present, the marker not only agrees with the subject argument but rather realises it directly. Only if a better candidate for subjecthood is present, will this be analysed as the subject, and the same verbal marker is then seen to function as an agreement affix:

It is generally assumed that, with the exception of French, Romance languages are so-called ‘pro-drop’-languages. As illustrated under (1) in a Spanish example, they are capable of realising the subject-argument through the morphological categories of Person and Number on the verb – categories which function as agreement markers as soon as a nominal phrase (NP) appears in the position of the subject.

(1) a. María, cantó [agreement]
   b. Cantó [subject realised by means of a verbal suffix]
   (translated from Kailuweit 2005:1).

This is intuitively appealing on the one hand, because it no longer leaves cantó (sing.3SG.PST.IND) without a subject as if it were ‘incomplete’ in any respect; on the other hand it is also problematic because it is unconvincing to treat one and the same affix alternatingly as an argument or as an agreement marker, depending on the presence or absence of a more suitable alternative. This solution has all the air of stopping short halfway. Once we have conceded that Spanish verbal affixes actually realise their subject arguments in some cases, we might then also consider the possibility that Spanish has come to be even more head-marking than that. Like all Romance languages, Spanish has a fully developed subject co-reference morphology on the verb and freely elidable lexical or pronominal subjects, a feature that has usually been described as pro-drop or null-subject. This, of course, only makes sense, if one conceives of Spanish subject co-reference in the same terms as of German subject agreement. If, however, the situation in Spanish were to be more analogous to Basque than to German, its subject-markers would have to be seen as an

13 French subject clitics which have traditionally been seen as clitical pronouns are nowadays analysed as agreement affixes by most linguists. In this case, French would be a null subject language, just like the rest of her Romance sister languages (cf. Kaiser 1992:115 et passim).
entirely different breed of affixes, namely obligatory cross-reference affixes which serve to express the subject argument. Du Bois’ concept of ‘mention’ almost imposes itself here, as the facts of this aspect of Spanish syntax resemble Sacapultec (and Basque) considerably more than those Indo-European languages on which the traditional analysis is based.

Spanish is therefore on the one hand similar to German in that it is not a null-subject language; both languages require overt subjects on the surface. The radical difference, on the other hand, lies in the morphological technique of subject realisation which is analytic in German and (poly-) synthetic in Spanish (and in the Romance languages in general). As would be expected for head-marking languages, lexical and pronominal ‘subjects’ are not integrated into the sentence core in the Romance languages and can either appear on the periphery (and only there!) or be elided altogether:

(19) [yo,i], a ese no se=lo=voy, a dar [., yo] (Spanish)
    I to that_one not him=it=go.1SG.PRS.IND to give I
(20) [moi,i] je=ne=vais, pas le=lui=donner [., moi] (French)
    I I=not= go.1SG.PRS.IND NEG it=him=give I
(21) [*ich,] dem gebe [ich] es nicht [*,ich] (German)
    I to that_one give.1SG.PRS.IND I it NEG I
    “I’m (certainly) not giving it to him!”

While subjects are realised through a head-marking construction in all Romance languages, the situation with objects is far less uniform across Romance. The general picture is complex and varies from language to language; they all seem to have in common, though, that at least certain objects are realised via head-marking on the verb. Spanish (and Catalan) are probably the two most advanced languages on the way towards clausal head-marking and I shall therefore use Spanish to illustrate the point (cf. Figure 7). That table is admittedly an extreme idealisation of the actual situation and is only intended to represent the gross overall picture. Thus, the only dependent-marking solution in the example is only possible as long as the direct object-noun phrase remains in its canonical position. Once it is raised into a left-dislocated topic position, the marker on the verb becomes obligatory:

(22)a. Juan conoc-e a María.
    John know-s to Mary
    “John knows Mary.”

b. A María, Juan la=conoc-e.
    to Mary John her.ACC=know-s
    “As for Mary, John knows her.”
The tendency in Spanish is clearly towards an increase of clausal head-marking and some varieties are already much further advanced in this than others. Kaiser (1992) reports:

In some varieties of Spanish, in particular those from the Rio de la Plata region in Argentina and Uruguay, the object clitic seems to be possible and even obligatory also in those cases where the co-referent direct object NP carries the semantic feature [+animate] and is therefore preceded by the preposition a (translated from Kaiser 1992:57).

In these dialects, Juan la conoce a María is completely grammaticised and we get the full head-marking situation!14

While it is already a noteworthy fact that Romance languages should exhibit clausal head-marking traits at all, the general typological importance of the shift depends heavily on the degree to which each language actually makes use of them. In an empirical study on the related topic of Romance object

---

14 Van Valin confirms this and acknowledges the existence of Latin American varieties of Spanish in which clitic-doubling has come to be obligatory (personal communication, 2002). Jaeggli (1982:14) reports that, in River Plate Spanish, clitic doubling is impossible with a non-pronominal, inanimate d.o., preferred with animate, specific d.o., highly preferred with a non-pronominal goal i.o., and obligatory in all other cases.
conjugation, Koch (1993, 1994) has undertaken the task of evaluating the statistic relevance of ‘clitic-doubling’ in French, Italian, and Spanish, the main results of which I have conflated into the synoptic table in Fig. 8. It shows a clear ranking, with French being most reluctant to adopt head-marking, Italian as a case in which both strategies seem to coexist on all levels, and Spanish with a clear-cut – and in the case of object pronouns already consummated – tendency towards clausal head-marking.

Further evidence in favour of a head-marking analysis of Romance objects comes from the classical examples of segmented sentences which have been particularly popular in the context of the debate over object conjugation in Romance and have quite justly (and following Queneau) led Koch (1993) to a comparison with Chinook and other Amerindian languages:

(23) a. [Il=la=lui=a=donnée], à Jean, son père
    he.NOM=her.ACC=him.DAT=has given to John his father
    sa moto (Tesnière 1959:175).
    POSS motorcycle.F.SG
    “His father gave John his motorcycle.”

   I in that house the feet NEG there=them=put.1SG.FUT never
   “I shall never set foot in that house.”

In languages without morphological case, the fundamental syntactic relations have to be encoded sequentially, with the ensuing consequence that the order of elements in the sentence becomes increasingly fixed and can no longer be used for discourse-pragmatic functions. While French exhibits a vast array of syntactic constructions to achieve this, of which only some exploit head-marking, these clearly seem to be the preferred structures in Spanish and Catalan. They provide a perfect solution to the discourse pragmatic problem of underlining the informational profile of an utterance by detaching lexical from grammatical information. Syntax proper is condensed into the morphosyntax of the clitic group, and the noun phrases, which carry the lexical information, are set free from the sequential restrictions of the core sentence and can now be used in the sentence periphery, where they can be ordered in any way the distribution of old and new information may require. Even if French shows certain head-marking tendencies as well, they have never come to be fully grammaticised as in Spanish.

15 These traditional examples are obviously and admittedly fabricated by linguists and rather unlikely to ever appear in spontaneous discourse, as they violate the Preferred Argument Structure constraint “Avoid more than one lexical core argument” (Du Bois 2003:34ff.). They are nevertheless grammatically possible utterances in these languages and illustrate their capability of representing all core arguments as affixes on the verb while converting the corresponding NPs into appositions outside the sentence core.
<table>
<thead>
<tr>
<th></th>
<th>French</th>
<th>Italian</th>
<th>Spanish</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct object</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VO = object follows verb; no clitic</td>
<td>92,0</td>
<td>85,5</td>
<td>[-hum] 95,0</td>
</tr>
<tr>
<td>OV = object precedes verb; no clitic</td>
<td>1,7</td>
<td>1,8</td>
<td>[-hum] 0,7</td>
</tr>
<tr>
<td>O,V = object follows verb; pleonastic clitic precedes verb</td>
<td>4,0</td>
<td>4,4</td>
<td>[-hum] 0,8</td>
</tr>
<tr>
<td>O,V = object precedes verb; pleonastic clitic precedes verb</td>
<td>3,0</td>
<td>6,4</td>
<td>[-hum] 3,6</td>
</tr>
<tr>
<td>Indirect object</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VI = object follows verb; no clitic</td>
<td>95,0</td>
<td>70,1</td>
<td>40,0</td>
</tr>
<tr>
<td>IV = object precedes verb; no clitic</td>
<td>0,0</td>
<td>2,4</td>
<td>0,0</td>
</tr>
<tr>
<td>I,V = object precedes verb; pleonastic clitic precedes verb</td>
<td>0,0</td>
<td>14,0</td>
<td>41,8</td>
</tr>
<tr>
<td>I,V = object precedes verb; pleonastic clitic precedes verb</td>
<td>5,0</td>
<td>8,7</td>
<td>18,2</td>
</tr>
</tbody>
</table>

All numbers are percentages. For details about the corpus used cf. Koch (1993:173). O = direct object, o = direct object clitic; I = indirect object, i = indirect object clitic; V = verb; Pron = emphatic, free pronoun; [+/- hum] = object with a human vs. non-human referent.

Figure 8: Percentage of pleonastic object clitics in Romance languages (data from Koch 1993)
This is mainly due to the early development of grammaticalised cleft-constructions in French which constituted an important alternative way of expressing theme-rheme structures, and while clefting never really caught on in Spanish, it eventually won out over clausal head-marking in French. The intermediate position of Italian as seen in the statistics of Fig. 8 is most probably due to the fact that clefting is more important in Italian than in Spanish but at the same time less important than in French.

6. Language change: intermediary stages of a change from dependent-marking towards head-marking

The last 1000 years have seen Spanish pronouns change from free forms similar to German pronouns towards increasingly bound forms. In a convincing article from 1990, Rini has tried to reconstruct and date the process in which Old Spanish clitic pronouns developed into the grammaticalised verbal affixes of Modern Spanish (Rini 1990). He does this by analysing the history of three related syntactic phenomena, namely 1. Interpolation, 2. Word order, and 3. Pronominal duplication.

Old Spanish clitics, though already phonologically dependent, had not yet developed into bound morphemes (cf. Rivero 1986:775). Their position with respect to their host was not yet as fixed as it is in Modern Spanish (e.g. él dixo, él lo dixo, lo él dixo “he said it” all can be found in the same document) and the so-called ‘interpolation’ of material between the clitic and its host was quite frequent:

(24) quien te algo prometiero (Corbacho 145)
    who pron.2SG.DAT something should_promise
    “whoever should promise you something”

    para lo mejor complir que lo ella non mando (Corbacho 91)
    for it better accomplish that it her not order.1SG.PRS.IND
    “in order to accomplish it better that I shall not order her to do it

While clitic-doubling did already occur in Old Spanish, it was far from being as frequent as today and it was never obligatory. Thus, clitic-doubled examples like:

(25) E tu as me tollido a mí un capellano (Berceo, 229d)
    and you.SG have me taken from me one priest
    “And you have taken away a priest from me.”

---

occur freely alongside non-doubled structures in the same text:

\[(26)\]  
\[
\text{A ti lo digo, nuera} \quad \text{(Corbacho ch. 38, p. 144)}^{18}
\]
\[
to \quad \text{you.SG it I say daughter-in-law}
\]
\[
\text{“I say it to you, daughter-in-law.”}
\]

But even where clitic-doubling did occur, as in (25), the construction is syntactically quite different from the corresponding structures in the modern language:

In Modern Spanish [...] the clitic is indispensable, which suggests that it is the basic pronoun and the tonic is the additional or redundant form, which appears for emphasis. There is no reason to believe, however, that in Old Spanish the emphatic constructions were of this same nature as regards the question of which was the basic and which was the redundant or additional element. In fact, the evidence provided by a diachronic perspective seems to suggest just the opposite: that the tonic pronoun, not the clitic, was the basic element of the emphatic constructions, and the clitic was additional or redundant when the two forms occurred together (Rini 1990:360).

What Rini found out was that, with respect to all three aforementioned parameters, Spanish seems to have reached the modern type of distribution at roughly the same time for each one of them, i.e. around the beginning of the 17th century.\footnote{Quoted from Rini (1990:360).} By then, no further examples of interpolation appear in the texts, pronominal clitics essentially become fixed to the left of finite and to the right of non-finite verbal forms and clitic-doubling becomes obligatory with emphatic pronouns. The fact that all three changes reach their conclusion more or less simultaneously is a strong indication that they have to be seen as three facets of one and the same grammaticalisation process which at this time must have reached a new stage: the stage of affixes.

Most interesting, however, is Rini’s observation regarding clitic-doubling. While in Old Spanish clitic-doubling the free pronouns acted as indispensable head of the construction and the clitics were just a redundant reinforcement, the respective roles can be shown to have changed completely in what Rini aptly dubs a ‘cephalic shift’ (Rini 1990:361). From the beginning of the 17th century on it is no longer the free pronouns but rather the clitic affixes that constitute the head of the construction while free pronouns may but need not be added for emphasis.

Put in the terms of Bybee (2005), the high token frequency of doubling clitics leads to an increased entrenchment of object clitics in the morpho-syntactic structure and ultimately “to the grammaticization of the new

\footnote{Other authors date the point for completed grammaticalisation much earlier; thus, for Barry, this state is reached from the thirteenth century onwards (cf. e.g. Barry 1987:219). For further empirical detail cf. Eberenz (2000:175-208).}
construction and the creation of grammatical morphemes and changes in constituency" (Bybee 2005:6). Object-clitics were first phonologically reduced in that they lost their capability of being stressed; from now on, stress could only be placed on the accompanying free pronoun or full NP. As verbs and clitics were used together ever more frequently, they came to be perceived as a unit in which the clitics were reinterpreted as obligatory grammatical morphemes. Because free object pronouns lost their obligatory status at about the same time, the new verbal markers began to function as cross-reference morphemes and were now treated as the actual realisations of objects.

In modern Spanish, (1) clitic doubling has become obligatory with any (emphatic) object pronoun, (2) is regularly used with indirect object NPs, and sometimes occurs even (3) with direct object NPs (cf. Fig. 8). In all these cases, the verbal affixes co-occur (and morphologically cross-reference with) their respective lexical or pronominal appositive phrases. With nominal objects the extent and obligatoriness of clitic-doubling seems to be in constant flux, and any researcher adducing examples must provide exact information as to which dialect or variety of Spanish is envisaged – all clear indications of a still ongoing process of change.

The cephalic shift created a new type of morphology in Spanish which would best be described as ‘cross-reference’. Obviously, there remain large areas in Spanish syntax where cross-reference is not (yet) obligatory or even possible. On the other hand, there are large areas (mainly with emphatic object pronouns) where it has already been fully grammaticalised. While Spanish may still be far from Basque in this respect, it is nevertheless clearly moving in the same direction.

Traditionally, grammaticalisation theory envisages the typical development of free (emphatic) pronouns as ‘emphatic personal pronoun > clitic pronoun > agreement affix’ (cf. e.g. Croft 2000:157). In the case of Spanish we can see that this is at least not the only possible development and that we are again faced with the type of shortsighted exclusion of languages with (clausal) head-marking as deplored by Nichols (1986:115f.; see quote above). While this type of grammaticalisation may lead to the establishment of a new type of agreement affixes, it may also, alternatively, create cross-reference affixes thus turning a language into one with clausal head-marking. For centuries, Spanish speakers have chosen head-marking strategies in order to resolve discourse-pragmatic problems of information structure and these preferred but still optional patterns are now turning into structural patterns of grammar in just the way Haspelmath (1999:193) envisages it.

As “grammars code best what speakers do most” (Du Bois 1987:811), Spanish is accepting ever more clausal head-marking as fixed, obligatory constructions into its syntax and we may be witnessing the emergence of a more and more consistent head-marking syntax. What has frequently been described as a change in which clitic pronouns are grammaticalised into
agreement-morphemes must then be seen as an even more radical change, not into agreement morphemes but into cross-reference affixes.

The difference between French and Spanish in this respect is more of degree than of principle, as all Romance languages possess at least some of these head-marking constructions. Clitic-doubled verbal complexes like il la lui a donné (or sp. se lo ha dado, cat. la hi ha donat etc., all meaning “he has given it to him”) are indeed quite analogous to the corresponding structures in Basque:

(27) \[\text{Bere } ai\text{-}a_{\text{it}}-k_{\text{i}} \quad \text{Jon-}i_{\text{h}} \quad \text{bere}\]
\[\text{POSS.3SG } \text{father-DET.ERG.SG } \text{Jon-DAT.SG } \text{POSS.3SG}\]
\[\text{moto-}a_{ji} \quad \text{eman } d_{ji-0_{\text{h}}-\Theta}, \]
\[\text{motorcycle-DET.ABS.SG } \text{give } \text{ABS.3SG-AUX} \text{have.PRS.IND-DAT.3SG-ERG.3SG}.\]
“His father gave John his motor-bike.”

The following examples show the possibilities for expressing differences in functional sentence perspective and the similarity of the French and the Basque solutions:

(28) a. Il=la=lui=a donnée, à Jean, son père, sa moto.
   he=it=him=has given to John his father his bike
   “His father gave John his motor-bike.”

(29) a. À Jean, il=la=lui=a donnée, son père, sa moto.
   to John he=it=him=has given his father his bike
   “His father gave John his motor-bike.”

(30) a. Sa moto, il=la=lui=a donnée, à Jean, son père.
   his bike he=it=him=has given to John his father
   “His father gave John his motor-bike.”

The main difference here lies in the fact that this cross-reference marking with appositional complements is the obligatory normal case in Basque while it is a stylistically marked exceptional construction, the appearance of which is bound to a discourse-pragmatic motivation. The French examples are all instances of segmented phrases, a discourse pragmatic repair strategy which justifies the apparition of the object clitics in French. With all lexical arguments in their conventional slots, clitic doubling becomes unacceptable in French:
(31) a. *Son père [*il=la=lui=] a donné sa moto à Jean.*
    his father he=it=him= has given his bike to John

b. *Bere aitak eman dio bere motoa Joni.*
    his father given he_has_it_to_him his bike to John
    “His father gave John his motor-bike.”

In Spanish, on the other hand, the corresponding sentence may not be the preferred one, but is at least grammatically possible:

c. *Su padre se=la=ha dado la moto a Juan.*
    his father him=it=has given the bike to John

On this basis we can tentatively formulate a typological hierarchy, where the Romance languages can be situated at an intermediate level:

<table>
<thead>
<tr>
<th>Type</th>
<th>example</th>
<th>Syntactic consequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head-marking</td>
<td>Abkhaz</td>
<td>– all arguments realised polysynthetically on the finite verb;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>– NPs are facultative appositions without any case marking;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>– no deictically unmarked object-pronouns (if at all, only demonstratives).</td>
</tr>
<tr>
<td>Double marking</td>
<td>Basque</td>
<td>– all arguments realised polysynthetically on the finite verb (head-marking);</td>
</tr>
<tr>
<td></td>
<td></td>
<td>– NPs are facultative appositions with case marking (additional, redundant dependent-marking);</td>
</tr>
<tr>
<td></td>
<td></td>
<td>– no deictically unmarked object-pronouns (if at all, only demonstratives).</td>
</tr>
<tr>
<td>Split marking</td>
<td>Spanish, mixed;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Catalan,</td>
<td>– Subject head-marking on the verb with appositional subject-NPs</td>
</tr>
<tr>
<td></td>
<td>European</td>
<td>or -pronouns respectively;</td>
</tr>
<tr>
<td></td>
<td>Portuguese</td>
<td>Objects are realised</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1. as affixes on the verb,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. as affixes + apposition and</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. as NPs (but not as pronouns!)</td>
</tr>
<tr>
<td>Dependent-</td>
<td>German</td>
<td>– all arguments realised as NPs pronouns with case marking;</td>
</tr>
<tr>
<td>marking</td>
<td></td>
<td>– only head-marking feature is the redundant subject agreement on the verb.</td>
</tr>
</tbody>
</table>

Figure 9: Head- vs. dependent-marking-hierarchy
7. **Conclusion**

Syntax, morphology, and phonology are usually so inextricably intertwined in diachrony that trying to establish exactly where a change originated may come down to the classical hen-egg problem. As Romance languages lost their case inflections (morphological phenomenon induced by phonological processes), syntactic relationships increasingly came to be encoded in the sequencing of constituents (syntax). As this sequencing became ever more rigid, new focusing techniques had to arise, as constituents could no longer be arranged freely to the discourse needs of language users (loss of word stress by the 16th century may have aggravated this problem in the case of French). The new techniques have in common that they place a pronominal element (relative or personal pronoun / clitic) in its canonical place such as to leave sequencing unaltered in the sentence core. The corresponding lexical NPs are thus freed to be placed in the periphery where they may be focused according to the conversational needs. The main solutions along these lines are clefts and pronominal reprise (or clitic-doubling), and while French and Spanish both use clefting as well as pronominal reprise techniques, the former are the typical solution in French, whereas Spanish largely depends on the latter.

As Duffer (this volume) has shown, the success of *c'est*-clefts in French let this constructions become so entrenched that their initial motivation as a focusing device didn’t remain their only function; with the advent of the ‘informative-presupposition cleft sentence’ in the 16th century, French clefts could now also carry given or thematic information. What Duffer calls clefting ‘beyond necessity’ can analogously also be observed for Spanish clitic-constructions, the history of which shows a continuous increase of uses that clearly go ‘beyond necessity’, and what once started out as individual speakers’ discourse strategies has, through the centuries, lost its dependence on the original motivations and has now been integrated into the very syntax of the simple sentence in Spanish. The typological drift from a conventional central IE language with agreement towards the ‘exotic’ type of clausal head-marking with cross-reference morphology can thus be argued to derive ultimately from new discourse strategies as a reaction to a syntactic change, namely the progressive fixation of word order towards rigid SVO (even if Langobardi’s Inertia Theory would have us expect otherwise).

The cross-reference analysis of Spanish clitics is not a mere terminological reformulation of the traditional analyses. The corresponding language type has long been well established and well described, although it has, until recently, received very little attention within IE linguistics and Romance philology in particular. If it can be shown that Spanish syntax is in a central aspect moving away from its IE sister languages like French, English, or German to become more like Basque, Abkhaz, or Lakhota, this is in itself an important typological insight. For the syntactic analysis of Spanish, the cross-reference hypothesis implies that, in many cases, lexical NPs as well as pronouns must no longer be
seen as the subject and object arguments, but rather as appositional phrases outside the sentence-core forming ‘mentions’ together with cross-reference markers on the verb. These mentions are discontinuous realisations of arguments in which the verbal markers are obligatory (and therefore the heads), while the lexical NPs or pronouns are syntactically facultative. Spanish does therefore neither tolerate null-subjects nor null-objects. A further consequence of this typological drift is that we can expect clitic-doubling to become ever more obligatory and to be carried over to further new domains.

8. **Closing remarks: no “object-conjugation” in Romance (or elsewhere)!

The phenomena discussed here are widely identical with those that have been invoked in the debate among Romance linguists over a supposed tendency towards ‘object-conjugation’ in these languages.²⁰ It should be clear from the above, however, that the notion of object-conjugation is misleading here as it evokes the notion of conjugation and with it the notion of agreement. As we have seen, agreement proper is only to be found in pure dependent-marking languages whereas the phenomena in question represent a typological shift precisely in the opposite direction, namely towards head-marking. While allusions to Basque and Hungarian are present in all contributions to the debate, the syntactic facts of these languages are consistently presented within a terminological framework developed for the description of dependent-marking Indo-European languages, treating Basque and Hungarian cross-reference elements as agreement-markers. This type of ‘agreement’ is then projected onto the Romance material. We must therefore reject the notion of ‘object-conjugation’ altogether, not only for the Romance phenomena, but also for the languages in which it purportedly is most fully developed, namely Basque and Hungarian. By failing to differentiate between agreement and cross-reference, the theory of Romance object-conjugation continues to look for the arguments where they are not to be found – in the noun phrases and strong pronouns – and perpetuates a strict distinction between a dependent agreement-marker and its nominal referent. In so doing they miss the very nature of this ongoing process of change: the emergence of cross-reference structures and clausal head-marking.

**REFERENCES**


