“Our first task in the study of a particular structure implicit to adult language behavior is to ascertain its source rather than immediately assuming that it is grammatically relevant.”

Bever 1970

ABSTRACT

Language research is particularly concerned with understanding the interplay between syntactic structure, variable aspects of linguistic form and language processing mechanisms. This paper discusses some recent results from Basque, an ergative language, that appear incompatible with the assumption that there are universal subject-gap advantages in relative clause processing. I argue that, given the incremental nature of language processing, the relative ease of processing structures involving subject or object gaps results from the combination of morphological specifications such as the ergative/nominative divide, and input-initial processing choices. This suggests that processing effects directly stated over notions like “subject of” or “object of”, widely considered derivative and not universal in Linguistics, are not sufficiently general. Abstracting away from these descriptive labels, and seeking alternative explanatory primitives that are not language-dependent is likely to yield a more parsimonious picture of processing complexity in language.

1. Introduction.

Research often reveals the abyssal depth of our ignorance in what might initially have seemed like a shallow puddle; discovering this is essential to formulate new questions, to find new answers to old ones and to discard those questions that turn out to be ill-formed. Carol Chomsky saw a sea of mystery in a child’s tentative, “half-cooked” language, though according to most linguists and thinkers of her time this pool of evidence held shallow scientific potential. However, her experimental approach to language acquisition revealed that syntactically complex structures took significantly longer to master than was originally assumed, and exposed the gulf between children’s language production and their linguistic knowledge. It also underscored the importance of finding new means to study language, not solely based on what speakers say, but also on how they comprehend what is said to them (C. Chomsky 1969).

Relative clauses are among those complex syntactic structures that children master relatively late, and adult language processing displays asymmetries that intriguingly parallel those found in acquisition. Children do not understand relatives in an adult-like
fashion before the age of six (Roth 1984), and they start producing them at about three years of age (Crain, McKee & Emiliani 1990). Relative clauses containing subject gaps (1a) are acquired and produced earlier than relatives containing object gaps (1b) (Brown 1972). It has been widely assumed in the literature that these asymmetries are invariant across languages and rooted in deep universal aspects of linguistic structure. I will contend instead that they are subject to variation, and depend on external aspects of linguistic form largely independent of syntactic structure, though extremely relevant to the study of language use.

I will discuss how certain aspects of linguistic form specific to a language belonging to an understudied type of languages (ergative languages) yield processing results and acquisition patterns that have hitherto hardly ever been reported from studies of a well-studied type of languages (nominative languages). Specifically, I will discuss some recent results from studies on relative clause processing in Basque that are incompatible with the widely-held assumption that subject-object language processing asymmetries are universal and that they tap into deep aspects of linguistic structure involving the core grammatical functions “subject-of” and “object-of”. As I will argue, the processing results obtained in Basque do not entail that the structural location of subjects and objects in ergative and nominative languages is different; rather they entail that morphological differences and input-initial choices have non-trivial consequences for processing.

2. Subject-Object asymmetries in language processing.

Studies from many languages report that relative clauses containing a gap in subject position are easier to process than relative clauses containing a gap in object position. These two types of relative clauses are illustrated in (1a) and (1b) for English.

(1)  a. The woman$_i$ [CP that $e_i$ saw the man] arrived early.
    b. The woman$_i$ [CP that the man saw $e_i$] arrived early.

This effect has been widely reported for languages where relative clauses follow their antecedent noun, such as English, Dutch, French, German and Spanish. In other languages, relative clauses precede the antecedent noun, and the gap position precedes its antecedent, as in the Japanese examples in (2), where (2a) is a subject-gap relative and (2b) is an object-gap relative.

(2)  a. [e$_i$ uma-o      ketta] robai-ga$_i$ sinda.
    horse-acc kicked mule-nom died
    “The mule that kicked the horse died.”

    b. [CP uma-ga   e$_i$ ketta] robai-ga$_i$ sinda.
    horse-nom kicked mule-nom died
    “The mule that the horse kicked died.” (from Ishizuka 2005)

---


Several hypotheses have been put forth to explain these subject-object asymmetries in relative clause processing. These accounts can be grouped in two classes: One group of hypotheses argues that this effect is rooted in fundamental properties of subjects and objects, and therefore predict the effect to be universal, revealing a fact of the internal architecture of human language. Another group of hypotheses argues that the effect is ultimately due to the temporal arrangement of the linguistic input, and thus predicts that the effect can be reversed if the temporal arrangement of the input varies.

Let us consider the first group of accounts. Universal subject-preference hypotheses argue that subject-object processing asymmetries follow from the greater saliency of subjects relative to objects in human language, where this saliency can be differently conceived, either as an inherent property of grammatical functions, understood as linguistic primitives, or as derived from general properties of syntactic structure.

The Accessibility Hierarchy (Keenan & Comrie, 1977) ascribes the advantage of subjects to an inherent cognitive preeminence of this grammatical function as compared to others. In its original form it states that grammatical functions are universally arranged in a hierarchy that determines their relative accessibility for relative clause formation, where subjects are higher than objects, and therefore more accessible. The hierarchy was originally postulated to account for relativization patterns in language typology and was later extended to other domains of grammar, language acquisition and processing (see Kwon, Polinsky & Kluender 2006 for a critical discussion of its explanatory force). Given this account, (1a) and (2a) are easier to process because they relativize subjects, whereas (1b) and (2b) are harder because they relativize objects.

(3) The Accessibility Hierarchy

Subject > Direct Object > Indirect Object > Oblique > Genitive > Object of Comparison

The Perspective Shift Hypothesis (Bever, 1970; MacWhinney, 1977, 1982; MacWhinney & Pleh, 1988), states that sentential subjects set the discourse-perspective, and that a processing event involving a perspective-shift is more costly than a processing event where perspective is kept constant. Processing a subject relative clause entails no perspective-shift, but object relative clauses induce a shift to a new subject in the embedded clause, thus generating a complexity effect. That is, (1a) and (2a) are easier to process because they involve the same subject with no perspective shift, whereas (1b) and (2b) are harder to process because they involve different subjects in the main and relative clause, with a shift in perspective.
The Structural Distance Hypothesis (O’Grady et al. 2003) appeals to the saliency of subjects in syntax; subjects are higher than objects in syntactic structure, and thus the structural distance between antecedent and gap is always greater for an object-gap structure, because the distance involved in a syntactic operation, calculated in terms of number of nodes crossed is always larger for object dependencies than for subject dependencies. It is widely agreed in Linguistics that all known languages share this hierarchical arrangement (Chomsky, 1965, 1995). If this is a universal property of language, then also according to the Structural Distance hypothesis, subject-gap relatives must be easier to process independently of the language observed.

Relativized Minimality (henceforth RM, Rizzi, 1990; Belletti and Rizzi, this volume) is a general account of locality effects in syntax where intervention of a possible antecedent has an impact on syntactic dependencies. RM views relative clause processing asymmetries as emerging from intervention effects on the resolution of the syntactic dependency between the gap and its antecedent (Friedmann et al. 2009). If an antecedent-like phrase structurally intervenes between the antecedent and the gap, it will increase processing difficulty. This account predicts that, given clauses involving similar subjects and objects (where similarity is dependent on the features of the DPs), object relative clauses are harder to process than subject relatives due to the structural intervention of the subject DP between the object gap and its antecedent, an intervention that does occur in the case of the subject relative, as illustrated in (4):

(4) a. DP₁ [CP...e₁...[VP...DP...]]

b. [CP...e₁...[VP...DP...]] DP

b. DP₁ [CP...[VP...DP...e₁...]

d. [CP...DP...[VP...e₁...]] DP

The schematic structural representations in (4a) and (4b) replicate the syntactic skeleton of the subject gap relative clauses (1a) and (2a) respectively. Regardless of temporal order, no DP structurally intervenes between the gap and the antecedent. The reverse is true of (4c) and (4d), representations of the object gap relatives (1b) and (2b) respectively, where the subject DP structurally intervenes between the antecedent and the object gap. Both the Structural Distance Hypothesis and RM share the idea that processing difficulty depends on syntactic structure and is independent from the temporal order of linguistic elements. However, they differ significantly on the predictions they make regarding the similarity/dissimilarity of the phrases involved in the sentences, which, according to RM crucially modulate the processing difficulty of the structure. Thus, according to the RM account, processing difficulty increases as the similarity of the antecedent and intervener increases (Friedmann et al. 2009, Adani et al. 2010). In this respect, this account relies both on invariable aspects of linguistic structure and specific features of the arguments involved in the structures.

Many studies on subject/object relative clause processing have shown that several factors related to the features of the DPs involved in the relative clauses modulate the relative processing difficulty of object relatives (Kidd et al. 2007). The processing difficulty of object relatives is reduced when the antecedent DP is inanimate (Mak et al., 2002, 2006; Traxler, Morris, & Seely, 2002; Traxler, Williams, Blozis, & Morris, 2005;
Weckerly & Kutas, 1999), and also when the clauses contain a pronominal subject or a proper noun subject (Gordon, Hendrick, & Johnson, 2001; Warren & Gibson, 2002, 2005). More recently, Adani et al. (2010) have shown that dissimilarities in grammatical features like number also diminish the relative difficulty of object relatives. Finally, another factor that has been shown to modulate the asymmetry between subject and object relative clauses is morphological case; in particular, relative clauses where the case of the gap and the case of the antecedent match are easier to process than clauses where the gap and the antecedent have different cases (Sauerland & Gibson, 1998; Ishizuka 2005). Therefore, all these factors must be taken into account and controlled for to reliably inquire into the impact of subject versus object gaps in relative clause processing asymmetries.

The group of hypotheses that claim the temporal arrangement of the linguistic input to be the locus of the asymmetry include working memory (Ford, 1983; Frazier & Fodor, 1978; Wanner & Maratsos, 1978), integration costs (Gibson, 1998, 2000; Hsiao & Gibson, 2003), syntactic strategies such as Active Filler Strategy and the Minimal Chain Principle (Clifton & Frazier, 1989; Frazier & Flores d’Arcais, 1989), the simultaneous influence of syntactic and non-syntactic information (MacDonald, Pearlmutter & Seidenberg, 1994; Trueswell, Tanenhaus & Kello, 1993), and differences in word-order canonicity (e.g., Bever, 1970; MacDonald & Christiansen, 2002; Mitchell, Cuetos, Corley, & Brysbaert, 1995; Tabor, Juliano, & Tanenhaus, 1997). For an extensive review of these proposals, see Traxler et al. (2002) and Hsiao & Gibson (2003).

There are also frequency-based approaches, according to which processing complexity emerges from competition between alternative structures partially activated during processing, favoring of the most frequent structure, and rendering less frequent structures harder to process. (Boland, 1997; MacDonald, 1994; MacDonald et al., 1994; McRae et al., 1998; Spivey-Knowlton & Sedivy, 1995; Trueswell et al., 1994; Gennari & MacDonald, 2008).

One relevant aspect of some of these accounts is the appeal to the temporal interval between the gap and the antecedent, with increasing temporal distance correlating with increasing complexity, as in the Dependency Locality Theory (DLT, Gibson, 1998, 2000). The DLT predicts inverse asymmetry-effects depending on the precedence relations: in prenominal relative clauses, the linear/temporal distance between the gap and the antecedent is greater in the subject-gap relative (5a) than in the object-gap relative (5b), regardless of the relative order of the verb and the object\(^2\). The reverse is the case in languages with postnominal relative clauses (5c,d):

\[(5)\quad \begin{align*}
    \text{a. [CP e_{subj} object/verb] antecedent} & \quad \text{prenominal subject-gap relative} \\
    \text{b. [CP subject e_{obj}/verb] antecedent} & \quad \text{prenominal object-gap relative} \\
    \text{c. antecedent [CP e_{subj} verb/object]} & \quad \text{postnominal subject-gap relative} \\
    \text{d. antecedent [CP subject verb/e_{obj}]} & \quad \text{postnominal object-gap relative}
\end{align*}\]

Therefore, a processing account based on the temporal interval between antecedent and gap such as the DLT predicts that whereas subject-gap relatives will be easier to process in languages with post-nominal relative clauses, object-gap relatives will be easier in languages with pre-nominal relative clauses. This prediction is met in the results

\(^2\) This is indicated by the slanted bar in the examples: object/verb means either object-verb or verb-object order.
reported in Hsiao & Gibson (2003) for Chinese, but subsequent studies in Chinese, Japanese and Korean that report a subject-gap advantage cast doubts with respect to the validity of the DLT as a general account of subject-object asymmetries (see references above, in particular Lin (2008) for a general discussion of this issue).

3. Subject/object prenominal relative clause asymmetries in an ergative grammar.

In Carreiras et al. (2010), we studied the processing of subject and object relatives in Basque; our results showed faster and easier processing for object-gap relative clauses as compared to subject-gap relative clauses. Basque has prenominal relative clauses, like Chinese, Japanese and Korean. But unlike all previously mentioned languages, including also those with postnominal relative clauses like (1), it is an ergative language. Ergative languages mark actor/undergoer core arguments of the verb in a way that is different from how nominative languages do it; this difference crucially involves the grammatical functions of subject and object. Hence, the study of processing asymmetries in an ergative grammar becomes particularly relevant in order to ascertain its cross-linguistic validity.

Approximately twenty five percent of the world’s languages exhibit ergativity in their grammars. Ergativity has been described in detail, and many ergative languages have been documented and studied until now, though the task of describing and understanding the phenomenon is far from complete (for overviews see Dixon 1994; Johns et al. 2006; Aldridge 2008; McGregor 2009). Despite the growing amount of research on ergativity in Linguistics, and despite the growing number of languages explored in language processing research in recent years, ergative languages have hardly been the object of processing studies so far.

The salient property of ergative languages is that the morphological marking of core verbal arguments diverges from that found in nominative languages. Nominative languages differentiate two classes of core arguments: (a) subjects (both transitive and intransitive) and (b) objects. Ergative languages also differentiate two main core argument types, but the classes are different: (a) one class consists of transitive subjects, and it is referred to as ergative, and (b) the other class consists of intransitive subjects and objects, and it is referred to as absolutive. The ergative morphological pattern is illustrated in (6) for Basque:

(6)  

a. emakume-a-k gizon-a ikusi du
    woman-D-erg man-D seen has
    ‘the woman has seen the man’

b. gizon-a etorri da
    man-D arrived is
    ‘the man arrived’

3 A full characterization of nominative and ergative marking systems is slightly more complex than the one just provided. Interested readers are referred to Dixon (1994), Johns, Massam & Ndayiragije (2006), Aldridge (2008), McGregor (2009) and references therein, where overviews of the phenomenon are provided in greater detail for a variety of languages.
The object of the transitive sentence (6a) and the intransitive subject in (46b) are alike (they belong to the absolutive class), whereas the transitive subject in (6a) is different (ergative). In ergative languages, labels like absolutive and ergative describe most conveniently the morphological marking of core arguments. As can be seen by comparing (6) to their English translations, these two labels do not readily correspond to subject and object as they stand in this language. This trait makes ergative languages particularly relevant to the study of phenomena that involve subject-object asymmetries, precisely because these two types of arguments find no direct equivalents in the morphology of ergative languages. As has been repeatedly observed “…ergativity raises a number of important problems for linguistic theory. (...) One such problem is the status and universality of subject (and to a lesser extent, object) as a grammatical relation, given the morphological groupings of ergative languages, (...) Ultimately we are led to questions of how grammatical relations are theorised.” (McGregor 2009:501)

The processing study in Carreiras et al. (2010) exploited a morphosyntactic ambiguity first used in Erdocia et al. (2009) for the exploration of word order processing complexity effects. This ambiguity involves the –ak ending on DPs4, illustrated in (7). The ending is homophonous with a singular ergative morphology and a plural absolutive morphology. In each of the cases, the morphological structure of the NP is different. Plurality in Basque is marked solely in the Determiner, which can be singular -a, as shown in (7a), or plural –ak, as shown in (7b)5. Determiner Phrases must be marked with ergative case when they are transitive subjects, and the form of the ergative case marker is –k. As shown in (7c), the result of adding ergative case to a singular DP yields the sequence –ak, homophonous to the plural determiner in (7b). For completeness, (7d) shows the resulting form of merging the plural determiner –ak with the ergative marker –k, which is the ending –ek, unambiguously denoting a plural ergative DP.

(7)  
|   | a. emakume-a | c. emakume-a-k |
|   | woman-D     | woman-D-erg    |
|   | “the woman” | “the woman (ergative case)” |
| b. emakume-ak | c. emakume-ek |
| woman-Dpl   | woman-Dpl+erg |
| “the women” | “the women (ergative case)” |

Hence, upon encountering as input a word like emakumeak, two possible interpretations are compatible with Basque grammar: interpreting it as a singular ergative meaning “the woman”, or interpreting it as a plural absolutive meaning “the women”. Remember that in ergative languages, subjects of transitive sentences are marked with ergative case, and objects are absolutive (8):

(8)  
**a + k:** [singular Determiner] + [ergative case]

---

4 I will be using the label DP, although it is less familiar to non-linguists, who are more familiar with the label NP. The label DP “Determiner Phrase” is more accurate for this type of nominal structures (Bernstein 2001).

5 I leave aside other elements pertaining to the Determiner category, such as demonstratives, which are not relevant to this discussion. The Determiner –a, -ak is not always definite, that is, it is not always translationally equivalent to English the (Laka 1993).
In Carreiras et al. (2010) we constructed relative clauses building upon this ambiguity. The experimental sentences were ambiguous between a subject-gap or an object-gap relative clause interpretation until the verb at the end of the main sentence was encountered (9):

(9) Object-gap relative clause

\[
\begin{align*}
\text{a. } & \text{[RC woman-D-erg [VP e_i \text{ seen} ] man-Dpl friend-Dpl are now} \\
& \text{[RC woman-D-erg [VP e, seen] has-C] man-Dpl friend-Dpl are now} \\
& \text{RC woman-D-erg [VP e, seen] has-C] man-Dpl friend-Dpl are now} \\
& \text{RC woman-D-erg [VP e, seen] has-C] man-Dpl friend-Dpl are now}
\end{align*}
\]

Subject-gap relative clause

\[
\begin{align*}
\text{b. } & \text{RC e_i [VP woman-Dpl seen] has-C] man-D-erg, friend-Dpl has now} \\
& \text{RC e_i [VP woman-Dpl seen] has-C] man-D-erg, friend-Dpl has now}
\end{align*}
\]

The sentences were presented word by word, and given the ambiguity of the word strings, the asymmetric cost should arise at the final main verb. Note that there are no differences between the two types of clauses in (9) in terms of storage resources (Gibson 2000), because the number of unresolved dependencies is the same in the two clauses: only one verb is required to generate a grammatical sentence\(^6\). The effect of word order canonicity can also be disregarded as a potential cause of an asymmetry (MacDonald & Christiansen, 2002), because in neither type of relative clause overt elements follow the SOV canonical order of the language (Erdocia et al. 2009).

In order to avoid effects due to the internal make-up of the arguments involved in the clauses, all the DPs in the experimental sentences are animate definite descriptions, and the cases of the gap and the antecedent always match: the gap and the antecedent bear absolutive case in the object relatives (9a), and the gap and the antecedent bear ergative case in the subject relatives (9b). Hence, the results reported here refer only to processing events were two animate definite descriptions are involved in the relative

\(^6\) A grammatical sentence is created by simply adding a verb: \text{ikasleak datoz “the students arrive” akin to examples (3b), (6c), using a synthetically inflected verb.}
clauses, and the gap and the antecedent are matched for case. The results do not preclude a modulation of processing difficulty determined by dissimilarities between the arguments of the clause or case-mismatch effects. Our study intended to control for these ulterior factors and did so by making all arguments involved similar regarding their lexical/internal and morphological/external structures.

Three experiments were carried out: two self-paced reading experiments and one ERP experiment. The only difference between the first and second self-paced reading experiment was the addition of a temporal adjunct after the verb of the main sentence in the second experiment, to avoid possible wrap-up effects at the end of the clause. The results of the two self-paced reading experiments showed that subject-gap sentences took significantly longer to read than object-gap sentences in the main verb (*dira/ditu*), the critical disambiguating region (10):

(10) Reading times for subject-gap versus object-gap relatives

(a) Experiment 1.

![Graph](image1)

(b) Experiment 2.

![Graph](image2)

In addition, an ERP experiment was conducted with the materials of the second experiment. ERPs showed that subject-gap sentences produced larger amplitudes than object-gap sentences in the P600 window immediately after reading the critical disambiguating word (*dira/ditu*), which also indicates greater processing difficulty for subject-gap sentences than for object-gap sentences.

Thus, the results of the three experiments conducted clearly point to a greater processing difficulty for subject-gap sentences than for object-gap sentences, contrary to the
majority of studies previously conducted in nominative languages. If subject-gap syntactic structures are not universally easier to process than object-gap structures, then accounts based on the inherent saliency or higher structural position of subjects cannot constitute a cross-linguistically valid account for processing asymmetries involving subjects and objects.

In order to determine whether frequency of occurrence correlated with these results, an initial corpus study was undertaken on a 25,000 word subset of the EPEC corpus (Aduriz et al., 2006), a 55,000 word sample collection of written standard Basque created and morphosyntactically tagged by the IXA group from the University of the Basque Country. Out of a total of 625 pre-nominal relative clauses found in the sample, 399 were subject-gap relatives (approximately 64%), while only 226 were object-relatives (approximately 36%). Subsequently, Yetano (2009) conducted a frequency study on the entirety of the EPEC corpus (300 000 words); results revealed that out of the total 1509 relative clauses in the corpus that involved subjects or objects as gaps, 65.6% were subject relatives, whereas 34.4% were object relatives. These data make a constraint satisfaction approach based on frequency not suitable to account for the findings: there is no correlation between the frequency of occurrence of subject versus object relative clauses and the processing asymmetry found. If frequency were the factor modulating processing difficulty, then the subject gap relative clause should have turned out to be easier to process than the object gap relative, contrary to results.

The object-gap relative clause advantage found in Basque is compatible with the Dependence Locality Theory (Gibson 1998), based on temporal/linear distance, because the object-gap in (9a) is temporally/linearly closer to its antecedent than the subject-gap in (9b). However, as mentioned before, results form other languages with prenominal relative clauses, where a subject preference has been reported, cast doubts on the cross-linguistic explanatory power of the DLT (Lin 2008). If results from languages with pre-nominal relative clauses do not converge, then some other factor or factors must be at play behind the inverse effect found in Basque.

If neither temporal nor structural distance can provide a comprehensive account of subject/object relative clause processing asymmetries across languages, we must seek alternative factors underlying the effect. Language specific properties are a plausible candidate, because language processing must handle externalized language forms, and infer syntactic structure there from. Given the view that the most plausible locus for language variation is morphology (Chomsky 1995), and given the fact that ergativity in Basque is a morphological phenomenon (Levin 1983), this linguistic trait stands out as a likely source for this divergent pattern of processing asymmetries, because it directly involves morphological case marking of core arguments.

For instance, if we consider morphological markedness, core argument marking in ergative languages entails that, given a transitive clause, the object is generally unmarked, and morphologically identical to the subject of an intransitive clause. For the case of Basque in particular, this is certainly borne out because the absolutive case carries no overt marker, whereas the ergative case is marked with the morpheme –k, as shown earlier in (6), (7) and (8).
If morphological markedness correlates with processing difficulty (Baayen et al. 1997; Badecker & Kuminiak 2007), then opposite patterns of complexity should arise in each type of grammar: nominative-accusative languages should typically display a nominative (=subject) advantage, but given the same underlying processing mechanism, ergative languages should display an absolutive advantage (remember that absolutive includes objects and intransitive subjects, that is, patient-like arguments but excludes transitive subjects, that is, agent-like arguments). Hence, the ergative morphological marking pattern can explain why an absolutive/object-gap advantage obtains when transitive relative clauses are processed in Basque.

An account involving morphological markedness and ergativity at its source gains plausibility given a recent relative clause comprehension study in Basque, which converges with the results we have just discussed. In a picture-matching task study, carried out two groups of children of 4 and 6 years of age respectively, and a group of young adults, Gutierrez (2010) found, for all three groups, that performance on absolutive/object-gap relatives was significantly better than in ergative/subject relatives: the former are comprehended with more accuracy than the ergative/subject relatives, which suggests a measure of processing difficulty.

These results are thus compatible with the hypothesis that morphologically unmarked antecedent-gap dependencies are easier to process: for nominative languages, unmarked dependencies correspond to nominative/subject-gap relative clauses, but in ergative languages unmarked dependencies are those established on absolutive arguments, which include objects. In the specific case of transitive sentences, as the ones considered here, this predicts a subject-gap relative clause advantage for the class of nominative languages, but an object-gap relative clause advantage for the class of ergative languages.

4. Subject processing strategies in an ergative grammar.

Another factor at play in the object-gap advantage found in Basque involves input-initial processing choices, which, given the incremental/deterministic nature of language processing, and the temporal order of the input, can lead to preferences that favour the object-gap interpretation in this language. Since it was first proposed by Frazier and Fodor (1978), it is generally accepted that when confronted with a morphosyntactically ambiguous structure, processing hypotheses favour the one leading to the simplest structure: “We assume in addition that if there is a choice of actions to take, then the processing system will again mirror the grammar, and so favor the ‘economy’ condition that the closest adjacent feature should be checked, rather than delaying to a later point in the derivation.” (Berwick, in press). This general principle of processing economy predicts that input-initial nominal phrases will be taken to be sentential subjects unless
there is overt evidence to the contrary. The prediction is strongly borne out across languages; there is ample evidence for a subject-first processing strategy for languages like Dutch, German, English, Italian, Turkish, some of which are head initial, others head final, some pro-drop and others not (e.g. de Vincenzi, 1991; Frazier, 1987; Schriefers et al. 1995, Demiral et al. 2008). It has been proposed that the subject-first preference is just a corollary of the minimalist nature of language processing, such that it postulates the minimal syntactic structure consistent with the input (Chomsky 1995, Bornkessel- Schlesewsky & Schlesewsky 2009, Berwick in press).

When encountering a morphologically ambiguous nominal constituent like \textit{emakumeak}, Basque speakers can in principle process that phrase as (i) a singular ergative case-marked DP “the woman” or as (ii) a plural absolutive DP “the women”, as illustrated earlier in (8). In turn, these two morphological choices combine with the fact that Basque is a pro-drop language where arguments need not be overt, and with the fact that word order is not fixed. Hence, upon encountering a phrase like \textit{emakumeak}, the grammar of the language allows at least either of the following (repeated here from examples in (8) for convenience):

\begin{align*}
\text{(12) a.} & \quad \textit{emakume-ak} \text{ etorri dira} \quad \text{Plural Intransitive Subject} \\
& \quad \text{women-Dpl arrived are} \\
& \quad \text{“The women have arrived”} \\
\text{b.} & \quad \textit{emakume-a-k} \text{ gizon-a ikusi du} \quad \text{Singular Transitive Subject} \\
& \quad \text{woman-D-erg man-D seen has} \\
& \quad \text{“the woman has seen the man”} \\
\text{c.} & \quad \text{pro}_i \textit{ emakume-ak} \text{ ikusi dit-u-zu}_i \quad \text{Plural Object} \\
& \quad \text{(you$_{erg}$) women-Dpl seen them-have-you} \\
& \quad \text{“You have seen the women”}
\end{align*}

Minimalist considerations rule out the choice in (12c), for it requires postulating additional syntactic structure to accommodate the null argument in a transitive structure. Indeed, minimal processing predicts that the choice for the ambiguous \textit{emakumeak} should be (12a), for this is the simplest syntactic structure consistent with the input; namely, the sentence containing an intransitive/unaccusative verb with a single argument. However, relative clause processing data indicate that subjects choose to interpret the ambiguous phrase \textit{emakumeak} as a singular ergative subject (12b), contrary to prediction. Let us see this in greater detail.

If the preferred choice for an ambiguous –\textit{ak} phrase were absolutive plural (12a) instead of ergative singular (12b), then incremental processing should generate a garden-path effect for the case of the object-gap relative, but no garden-path effect should emerge for the subject-gap relative. Consider (13a), where a sentence-initial DP like \textit{emakumeak} is encountered. Given that it can be an intransitive/unaccusative absolutive subject, minimal processing makes the choice of parsing it as a plural absolutive phrase “the women”. Once this choice is made, when the verb \textit{ikusi} “to see” is encountered (13b), we must now assume that \textit{emakumeak} is the object of a transitive sentence; given that Basque is a head-final, free word-order and pro-drop language, the transitive subject can appear later in the sentence or not at all. Notice that in an ergative grammar it is not an obvious matter whether the change from absolutive subject to absolutive
object involves a significant change in argument role-identification, of the kind that would be involved for a similar change in a nominative language (nominative to accusative). In any event, once the verb *ikusi* is encountered, *emakumeak* (if absolutive) must be the object of the transitive clause.

(13)  
\[
\begin{align*}
\text{a. emakume-ak} & \ldots \\
& \text{woman-D}{_\text{pl}} \\
\text{b. emakume-ak ikusi} & \ldots \\
& \text{woman-D}{_\text{pl}} \text{ seen} \\
\text{c. emakume-ak ikusi ditu-en} & \ldots \\
& \text{woman-D}{_\text{pl}} \text{ seen has-Comp}
\end{align*}
\]

When the inflected auxiliary verb is encountered (to which the complementizer is attached), given its form it becomes clear that (i) the ergative subject is singular and object agreement is plural (morpheme -it-) and (ii) this is a relative clause. Given the initial choice for *emakumeak* as a plural absolutive, and since the relative clause is now “closed” by the Complementizer –en, we must postulate the ergative gap in the clause, which is co-referent with the head of the relative immediately following the inflected auxiliary:

(14)  
\[
\begin{align*}
[\text{RC } e_i [\text{VP woman-D}{_\text{pl}} \text{ seen}] \text{ has-C}] & \text{ man-D}{_\text{erg}}, \text{ friend-D}{_\text{pl}} \text{ has now} \\
[\text{RC } e_i [\text{VP emakume-ak ikusi}] \text{ ditu-en}] & \text{ gizon-ak ditu orain} \\
& \text{“The man, that } e_i \text{ saw the women} \text{ has friends now”}
\end{align*}
\]

In other words, if subjects made an intransitive-subject choice for the ambiguous sentence-initial DP *emakumeak*, this should result in a subject-gap relative clause processing choice, contrary to the results obtained. Instead of this, speakers preferred to process the ambiguous sentence-initial DP as a singular ergative phrase, that is, as a transitive subject, which explains why object-gap relative clauses were processed faster and with greater ease:

(15)  
\[
\begin{align*}
[\text{RC emakume-a-k [VP e_i ikusi]} \text{ ditu-en}] & \text{ gizon-ak, lagun-ak dira orain} \\
[\text{RC woman-D}{_\text{erg}} [\text{VP e_i seen}] \text{ has-C}] & \text{ man-D}{_\text{pl}}, \text{ friend-D}{_\text{pl}} \text{ are now} \\
& \text{“The men, that the woman saw } e_i \text{ are friends now”}
\end{align*}
\]

Hence, a sentence-initial choice favouring the ergative over the absolutive has as a consequence that the subject-gap disambiguation at sentence final position results in a garden-path, while no such effect emerges in the case of the object-gap disambiguation. The increase in reading times for the subject-gap disambiguation as compared to the object-gap disambiguation shown in (10) strongly suggests this is in fact the case.

Given that ergative marked arguments are unequivocally subjects, this ergative-preference can be interpreted as strong evidence of a “subject-first” processing strategy, independently reported in most processing studies to date. However, it is

---

7 For reasons of space, I will not retrace in detail the consequences of an initial ergative choice regarding the object-gap preference; the interested reader can verify that repeating the previous steps illustrating the consequences of an initial absolutive choice replacing an initial ergative do yield an object-gap preference.
worthwhile noting that this processing choice in favour of the ergative would violate the principle of minimal structure (Chomsky 1995, Bornkessel-Schlesewsky & Schlesewsky 2009, Berwick in press), because minimalist processing predicts that in the face of ambiguity, the parser chooses the option that generates the simplest structure. If, as it is generally assumed, the simplest sentence structure is that corresponding to a monoargumental/unaccusative predicate, then we should expect that for the case of *emakumeak*, speakers should choose an absolutive interpretation, given that this is the form a subject takes in unaccusative predicates in ergative languages.

We must not conclude, however, that Basque violates the principles of minimal parsing. As mentioned before, all DPs in the experimental sentences in Carreiras et al. (2010) were *animate* definite descriptions. Ferreira & Clifton (1986) demonstrated that initial DP-animacy strongly influenced processing choices and garden-path effects; since then, the effects of animacy in relative-clause processing have been explored in depth (Traxler et al. 2002, 2005), and it has been suggested that animacy strongly determines the processing choices speakers initially make for relative clauses (Mak et al. 2002, 2006). Bornkessel-Schlesewsky & Schlesewsky (2008, 2009) argue that prominence features like animacy, known to be active features in the morphology of many human grammars, can drive processing choices in the absence of other cues, and can have variable impact on processing cross-linguistically. In particular, verb-final grammars would rely more on prominence features like animacy to determine actor/agent-like and undergoer/patient-like roles during online comprehension. The finding that morphologically ambiguous animate DPs in Basque are processed as ergative/transitive subjects appears to converge with the findings in Choudhary et al. (2007), where ERPs revealed N400 effects for clause-initial inanimate ergatives, that the authors interpreted as an index of difficulties in grammatical role assignment. In sum, the subject-initial processing strategy that follows from a minimalist processing is modulated by grammatically active features like animacy, which favour an ergative/actor processing choice for a sentence with an initial animate ambiguous form.

5. Subject/object postnominal relative clause asymmetries in an ergative grammar.

If morphological markedness combined with sentence-initial processing choices are the main factors behind the object-gap advantage found for relative clauses in Carreiras et al. (2010), then we predict that variations in the temporal arrangement of the linguistic input can have an impact on this effect. In Yetano et al. (2010) we explored this possibility by investigating a type of relative clauses in Basque, which is crucially postnominal. This type of relative clause is less frequent, and belongs to a higher register of the language. Unlike prenominal relative clauses, it involves a Wh-element in Complementizer position, much like an English relative clause. Examples of the postnominal relative clauses used in Yetano et al. (2010) are given in (16):

(16) Post-nominal subject-gap relative clause

\[
\text{a. gizon-a-k, } [\text{zeinak, } e_i \text{emakumeak ikusi bait-ditu}] \text{ lagunak ditu orain}
\text{man-D-erg, } [\text{who, } e_i \text{woman-D_pl seen C-has}] \text{ friend- D_pl has now}
\text{“the man who saw the women has friends”}
\]
Postnominal object-gap relative clause

b. gizon-ak, i [zeinak, emakume-a-k e, ikusi bait-ditu] lagunak dira orain man-Dpl [who, woman-D-erg e, seen C-has] friend-Dpl are now “the men who the woman saw are friends”

As shown in (16), we employed morphologically ambiguous –ak phrases to construct our experimental sentences. Hence, as in Carreiras et al. (2010), relative clauses were ambiguous between as subject-gap or an object-gap reading until the inflected verb of the main clause was reached (dira/ditu), in the next-to-last position. We generated 26 experimental sentences, 13 subject relatives and 13 object relatives, which were mixed with 74 fillers. The experiment consisted of phrase-by-phrase (=word-by-word) self-paced reading with a comprehension question after each sentence. Results from the 40 native speakers who took part in the experiment revealed shorter reading times for the subject-gap relatives at the critical disambiguating region and at the subsequent region, as shown in (17), and comprehension accuracy was higher for subject relatives (86% correct responses for subject relatives versus 81% correct responses for object relatives.

![Figure 1: reading times for the different regions of the sentences containing SR and OR clauses.](image)

The object-gap preference found in prenominal relative clauses in Carreiras et al (2010) is reversed in postnominal relatives, which display a subject-gap advantage. These results are compatible with a linear distance account like the Dependency Locality Theory (DLT, Gibson, 1998, 2000), because it predicts inverse asymmetry-effects depending on the precedence relations, of the type we find in Basque. However, studies from languages with pronominal relatives that report a subject-gap advantage render the DLT problematic as a cross-linguistic account of subject-object asymmetries (see Lin 2008 for a general discussion).
The results are also compatible with the hypothesis that the main factors behind the processing asymmetries are morphological markedness and sentence-initial processing choices. An interaction between an absolutive-gap advantage and subject preference that is parallel to the one discussed here for Basque is also recently reported by Polinsky, Gomez-Gallo, Kravtchenko & Testelets (2012) for relative clause processing in Avar, an ergative language from the Caucasus. In the case of Basque, ergative morphology (which has absolutive as the unmarked class), combined with a DP sentence-initial processing choice that favors an ergative interpretation for ambiguous animate DPs, yields the subject-advantage in the postnominal case, as can be seen by considering the syntactic structure in (16a), which takes the initial DP to be ergative, and thus leads to a subject gap interpretation, opposite to what was the case with the pronominal relative in (15). For convenience, we repeat the two representations in (18), where (18a) corresponds to a postnominal subject-gap relative, and (18b) to a pronominal object-relative:

(18) a. gizon-a-k [zeinak, e, emakumeak ikusi bait-ditu] lagunak ditu orain  
man-D-erg [who, e, woman-Dpl seen C-has] friend-Dpl has now  
"the man who saw the women has friends"

b. [RC emakume-a-k [VP e, ikusi] ditu-en] gizon-ak_i lagun-ak dira orain  
[RC woman-D-erg [VP e, seen] has-C] man-Dpl friend-Dpl are now  
"The men (that the woman saw e,) are friends now"

Both syntactic structures, different as they are, share the fact that the first DP encountered is interpreted as an ergative singular subject. The object-advantage of the pronominal relative and the subject-advantage of the postnominal relative both follow given this initial processing choice.

6. Conclusion: subject/object asymmetries need not be about subjects and objects.

The results and processing mechanisms we have discussed in this paper underscore that a full picture of language and the way in which it is processed in real time requires the study of a broad sample of significantly different languages and linguistic phenomena; only a truly cross-linguistic approach to language research will reveal the interplay of the various factors at work in molding the interface between grammatical form and language processing strategies.

The task of language processing is to infer syntactic structure from the temporal input in order to access (propositional) meaning. It is well established that several non-syntactic factors modulate the strategies involved in processing. Given this, complexity effects observed in processing do not necessarily reveal syntactic/grammatical complexity per se. Differential processing costs can result from the interplay of processing strategies that determine parsing choices for incoming input, as input is turned into language (words, phrases, sentences). A cross-linguistic outlook is thus essential to discover the ultimate nature not only of linguistic structure (as is widely accepted in Linguistics), but also of input-processing mechanisms at an adequate level of abstraction.

In generative linguistics, notions like “subject” or “object” are not taken to be universal primitives; rather, they are considered derivative-descriptive categories that label syntactic configurations (Chomsky 1965). It is plausible that also in language
processing, when encountering what appear to be irreducible subject/object asymmetries, these can be reduced to more general, and less language-dependent processing mechanisms. In this respect, the goals originally set out for linguistics in *Syntactic Structures* can be used to state what the ultimate goal of cross-linguistic studies is, not only in theoretical inquiry, but also when inquiring into language use:

More generally, linguists must be concerned with the problem of determining the fundamental underlying properties of successful grammars. The ultimate outcome of these investigations should be a theory of linguistic structure in which the descriptive devices utilized in particular grammars are presented and studied abstractly, with no specific reference to particular languages. (N. Chomsky 1957:11)

This goal can be extended to research on language processing, by simply writing “language researchers” instead of “linguists” then “processing mechanisms”, instead of “successful grammars”, and finally “language processing” instead of “linguistic structure”, thus:

More generally, language researchers must be concerned with the problem of determining the fundamental underlying properties of processing mechanisms. The ultimate outcome of these investigations should be a theory of language processing in which the descriptive devices utilized in particular grammars are presented and studied abstractly, with no specific reference to particular languages.

AKNOWLEDGEMENTS: I wish to thank Robert Berwick and Massimo Piattelli-Palmarini for the invitation to take part in *Rich Languages from Poor Inputs. A Workshop in Honor of Carol Chomsky*. I also thank the audience and participants, particularly Adriana Belletti, discussant of my talk. Thanks to Jon Andoni Duñabeitia, Manuel Carreiras, and Maria Polinsky, and to colleagues in *The Bilingual Mind* research group, Kepa Erdocia, Irene de la Cruz-Pavía, Mikel Santesteban, Iraia Yetano and Adam Zawiszewski, for their feedback and suggestions. Errors and misinterpretations are solely mine. This research has been funded by the Spanish Ministry of Education and Science (CSD2007-00012), the Spanish Ministry of Science and Innovation (FFI2009-09695), and the Basque Council for Education, Universities and Research (IT414-10).

REFERENCES:
Aldridge, E. (2008) Generative Approaches to Ergativity *Language and Linguistics Compass* 2/5, 966–995,


