Animacy in the adjunction of Spanish RCs to complex NPs
Carlos Acuña-Fariña a; Isabel Fraga a; Javier García-Orza b; Ana Piñeiro a
a Universidade de Santiago de Compostela, Coruña, Spain b Universidad de Málaga, Malaga, Spain

First published on: 01 March 2009
Animacy in the adjunction of Spanish RCs to complex NPs

Carlos Acuña-Fariña and Isabel Fraga
Universidade de Santiago de Compostela, Coruña, Spain

Javier García-Orza
Universidad de Málaga, Malaga, Spain

Ana Piñeiro
Universidade de Santiago de Compostela, Coruña, Spain

The present paper focuses on the role of animacy in the processing of relative clauses (RCs) after complex NPs. We follow research by the Desmet et al. team on Dutch in exploring the role of animacy in Spanish RCs. We present data from a corpus study and two self-paced experiments and we compare the three studies and the Dutch and Spanish results. Our main objective is to fill important gaps in past research on the processing of adjunction ties in Spanish and to offer a more detailed exploration of grain effects in exposure-based accounts. In particular, we have sought both to analyse the match between corpus studies and online processing in Spanish much more closely than it has been until now and to see whether animacy could revert the well-established tendency of Spanish RCs to attach high inside the complex noun phrase.

Keywords: Adjunction; Animacy; Relative clause; Tuning.

Ever since Cuetos and Mitchell (1988), the ambiguous construction exemplified in (1):

Someone shot the servant of the actress who was on the balcony (1)

Correspondence should be addressed to Carlos Acuña-Fariña, Departamento de Inglés, Facultad de Filología, Avda. Castelao s/n Santiago de Compostela, E-15782, Coruña, Spain. E-mail: iacanoe@usc.es

Preparation of this paper was supported by grants SEJ2005-08911/PSIC (the Spanish Ministry of Science), PGIDIT06PXIC204102PN, and INCITE07PXI211031ES (the last two from the Regional Autonomous Government of Galicia). Ana Piñeiro was the recipient of a predoctoral grant from the Galician Government. We thank Marc Brysbaert, Don Mitchell, Manuel Perea, and an anonymous reviewer for helpful comments on an earlier draft.

© 2009 Psychology Press, an imprint of the Taylor & Francis Group, an Informa business
http://www.psypress.com/ecp DOI: 10.1080/09541440802622824
has posed a significant puzzle to the language processing community of researchers. (1) features a complex noun phrase (CNP) which is followed by a modifying relative clause (RC). Its ambiguity resides in the fact that the modifying clause can be “related to” any one of the NPs forming the previous CNP. That is, in (1) both *actress* and *servant* may be on the balcony. Until 1988, it was customary to assume that locality, or recency, should determine the preferred resolution of any ambiguity involving two or more equally complex potential structures. For a structure like that in (1), that means that, given the competition between a distant (*servant*) and a local (*actress*) site, the RC would be expected to opt for the local one. There were two fundamental reasons for endorsing that view: (a) the sound logic of computational economy; and (b) the fact that English, the particular language most typically studied and thus more often implicitly taken to represent language in general, appeared to show an N2 bias indeed.

By showing that Spanish language users generally opt for an N1 bias instead, Cuetos and Mitchell (1988) seriously undermined the strength of those two seemingly reasonable factors (and the rationale for a universal parser), a situation that sparked two complementary lines of research. On the one hand, a host of new theories came to light as a result of the need to either replace or qualify the role of locality in at least [CNP + RCs]. On the other hand, a host of different languages were subjected to empirical investigation with a view to placing the Spanish–English disparate results in a wider context.

Starting with the latter line of research, the cross-linguistic one, most studies at least initially sided with the Spanish results. Thus, apart from Spanish itself (Carreiras, 1992; Carreiras & Clifton, 1993, 1999; Cuetos, Mitchell, & Corley, 1996), at least French (Frenck-Mestre, 1997; Zagar, Pynote, & Ratjeau, 1997), German (Hemforth, Konieczny, Scheepers, & Strube, 1998), and Dutch (Brysbaert & Mitchell, 1996) showed an N1 bias. English, however, presented a somewhat less consistent pattern of results, although the general tendency seemed to be a slight preference for the second site (Carreiras & Clifton, 1999; Corley, 1996; Fernández, 2000; Frazier & Clifton, 1996; Henstra, 1996). The disparity is particularly interesting from a linguistic point of view because it cuts across two language families: German and Dutch are Germanic languages which nevertheless apparently behave unlike English, another Germanic language, and more like Spanish, a Romance language. However, that is the least troubling of the circumstances affecting the [CNP + RC] debate. Soon, the English N2 bias was questioned by results from American English (Clifton, 1988). Later, Carreiras and Clifton (1993) found no reliable pattern of adjunction in either direction in English either. According to de Vincenci and Job (1993), Italian, a language of remarkable similarity to Spanish, shows an N2 bias, but this was later attributed to segmentation effects.
Carreiras & Clifton, 1993). Baccino, de Vincenci, and Job (2000) have however insisted that their own experiments prove Italian to be an N2 language, but again Frenck-Mestre and Pynte (2000b) have reasoned that such results are biased by the authors not controlling the internal structure of the CNP, particularly the role that different prepositions play in adjunction in the same CNP domain. Additionally, after the Frenck-Mestre (1997) and Zagar et al. (1997) experiments, in which French exhibited a robust N1 bias, Baccino et al. claim to have found evidence of an N2 bias instead, but this is also questioned by Frenck-Mestre and Pynte (2000b) on grounds of the internal structure of the CNP too. To complicate things even further, Fernández (2000) has found that monolinguals tested at the Complutense University in Madrid displayed an N2 preference in Spanish. Although the evidence for an N2 bias has been more consistently detected and controlled in experiments on the English language, some more recent evidence has suggested a similar bias in Arabic, Romanian, Swedish, Norwegian, and European Portuguese (Abdelghany & Fodor, 1999; Ehrlich et al., 1999; Maia, Costa, Fernández, & Lourenço-Gomes, 2004; Maia, Fernández, Costa, & Lourenço-Gomes, 2006). In sum, results are far from clear, although the general consensus seems to be that evidence is stronger for the higher site (see Carreiras & Clifton, 1999; Frazier & Clifton, 1996; Mitchell, Brysbaert, Grondelaers, & Swanepoel, 2000; and Cuetos et al., 1996, for reviews; see also Swets, Desmet, Hambrick, & Ferreira, 2007, on recent evidence that suggests that, when individual working memory is controlled for, English and Dutch readers react to attachment choices in a similar fashion).

As for the theories which have been proposed to account for [CNP + RCs], these reflect both in number and in nature what is now seen as its extreme complexity. So far, however, none has been able to account for all the findings since these are contradictory at various levels (different findings for the same language, for the same methodological technique used, and even for the same translations). Construal (Frazier & Clifton, 1996; Gilboy, Sopena, Clifton, & Frazier, 1995) had until recently steered clear of serious refutation of its principles. In construal terms, recency (or late closure) is exonerated from applying to nonprimary constituents of sentences, like RCs. The way in which construal aims to account for cross-linguistic variation rests therefore on two prongs: (a) Variation simply does not apply to primaries, which must therefore still obey late closure and other similar structural principles without exception in all languages of the world; and (b) variation may apply to dispensable constituents, which, due to their very dispensability, are processed relatively late, that is, in nonfirst-pass measures. Invoking late analysis within a Chomskian-based theta-domain (Chomsky, 1981), Construal theoreticians leave room in their model for all kinds of nonsyntactic forces affecting the eventual “association” of the RC to its
host. It is assumed that adherence or sensitivity to such forces (which include
semantic, pragmatic, collocational, and lexical biases) may vary cross-
linguistically. Adherence or sensitivity to syntax is not supposed to vary
cross-linguistically. The different assumptions belie a form of serial thinking:
first must come core syntax every time, everywhere, in all languages of the
world; then may come all the rest, including (presumably very late-acting)
Gricean maxims, any time and anywhere, as each particular language
permits.

Indeed, ever since Gilboy et al. (1995) showed that the internal structure
of the CNP, and most importantly, the theta-marking properties of the
intervening preposition, was essential in the determination of adjunction
preferences, the model’s credibility won considerable support. But problems
appeared soon. First, Zagar et al. (1997) did not find in French any trace of a
late-acting interpretative processor but, instead, a seemingly solid structural
bias towards the N1 site, regardless of context manipulation. Pynte and
Colonna (2001) have found that some other modifiers at least do not wait for
late interpretative information either, which means that the primary versus
nonprimary distinction that is at the heart of the model may not be all that
counts in the processing of adjunction. Besides, how precisely adjunction
remains in suspension (Deevy, 2000) and is later fully determined by a
myriad of late-acting factors, and how these are to be respectively ranked,
has never been clear (Mitchell et al., 2000; Traxler, Pickering, & Clifton,
1998). Nor is it clear how precisely the model can account for individual
differences and habituation patterns reported in the literature (Brysbaert &
Mitchell, 1996; Corley, 1996; Corley & Caldwell, 1996; Frenck-Mestre &
Pynte, 2000a; Garcia-Orza, 2001; Swets et al., 2007). Finally, Mitchell et al.
(2000) have shown that the Gricean explanation that proponents of the
model have put forward to account for different adjunction patterns in
[CNP + RCs] between English and most other languages examined (namely,
the availability in the former of other syntactic alternatives like the Saxon
Genitive) does not explain the facts of Dutch, or of Afrikaans, where the
Saxon Genitive is also used and yet there is a clear N1 bias, unlike English.

Another model that deserves mention here is Fodor’s prosody-based
“same-size-sister”. According to Fodor (1998), what differentiates results in
the [CNP + RC] construction from results obtained for other structures
where recency has proved to be strong is the size of the attachee relative to
that of the host. Relative clauses tend to be large modifying structures by
comparison with ordinary adjectival or prepositional phrase modification.
Assuming that a constituent “likes to have a sister of its own size” (p. 285), it
follows that light constituents will prefer to attach low, whereas heavy ones
will opt for high heads (or, to be more precise, for the combination of the
two nouns, the head of which is the higher one). As Fodor herself points out,
this is a peculiar antigravity law that rests on an implicit prosody hypothesis
(Fodor, 2002). Indeed, some studies (Fernández & Bradley, 2000; Igoa, 1999; Lovric, Bradley, & Fodor, 2000) have found size effects in the sense that a short RC like who left is preferably associated with the lower site in a sentence like Somebody shot the servant of the actress who left. A principle of balance would prevent the short RC from attaching to the long CNP (the servant of the actress). Fodor points out that the universal parsing strategies advocated by formalists can still be maintained if we assume that prosodic processing works in parallel with syntactic processing and has an impact on ambiguity resolution. Since languages differ in their prosodic packaging, cross-linguistic variation should in principle be easily accounted for. This proposal is in part a reformulation of Frazier and Fodor’s (1978) old Sausage Machine model in the light of the arguments put forward by Gilboy and Sopena (1996; on “prosodic visibility” and the syntax-prosody mapping, see Carlson, Clifton, & Frazier, 2001; Pynte & Prieur, 1996; see also Schafer et al., 2000, for similar results both in production and comprehension; for neurocognitive evidence of immediate use of prosodic cues in processing, see Steinhauer, Alter, & Friederici, 1999). An interesting aspect of this kind of research is that the “size issue” interacts with recent work on memory which shows that low span readers have a greater tendency to break up large segments of text because of their limited working memory, leading to a higher attachment of the RC (Swets et al., 2007). Again, the model seems to have had success in the work cited, but also problems (see Colonna & Pynte, 2001), although it seems fair to say that it is still being tested.

Indeed, at present only a review article can do justice to the complexity of the picture that emerges from experimental research on [CNP + RCs]. We suggest that, as a working hypothesis at least, all the nuances that have come out in over a decade and a half of research (recency, theta-marking, prosody, context, etc.), be seen as processing parameters whose exact timing and role in the whole process of understanding is yet to be determined. And we further suggest that just as the Gilboy et al. (1995) questionnaire study opened up a door for subsequent fruitful research on aspects (nuances) of theta-theory, it is now Desmet et al.’s series of studies on Dutch (Desmet, Brysbaert, & De Baecke, 2002; Desmet, De Baecke, Drieghe, Brysbaert, & Vonk, 2006) that may be showing us the way by bringing to light stark differences and interrelationships among potential processing parameters like animacy, modifiability or concreteness, none of which have a syntactic nature.

Desmet et al.’s work grew out of research aimed at defending the Tuning model (Brysbaert & Mitchell, 1996; Corley, 1996; Corley & Caldwell, 1996; Mitchell & Cuetos, 1991; Mitchell, Cuetos, & Corley, 1992; Mitchell, Cuetos, Corley, & Brysbaert, 1995). Based on the important role played by frequency in shaping other aspects of cognition, including facets of language (Duffy,
Morris, & Rayner, 1988), Tuning advocates claimed that the large picture of syntactic processing is also governed by statistical tallies. Thus, in the face of a syntactic ambiguity, the processor will simply select the alternative which has proved most frequently reliable in the past (Mitchell et al., 1995, p. 470). A crucial feature of the model is that the precise delimitation of the grain of analysis must be stipulated, as there are several concurrent levels of analysis over which frequency counts may act: the lexical, the syntactic, etc. (the so-called grain size problem; see Clifton, 1994; also Christiansen & MacDonald, 2000). A syntactic approach to Tuning postulates holds that the tallies that count are only those registering syntactic structures as wholes, and not just words. However, lethal evidence against this view came out when Mitchell and Brysbaert (1998) showed that, when such a coarse grain is examined, corpora studies and online measures do not match in Dutch (overall N1 preference online, N2 preference in corpus). Contradictions between corpora studies and online studies have also been reported in Gibson and Schütze (1999) with CNPs containing three, as opposed to just two, attachment sites (as in the lamp near the house with the door that...; see also Desmet & Gibson, 2003). Implications that production and comprehension might proceed on different rails could not be avoided.

Desmet et al. (2002) reexamined the Dutch data (especially the 1998 corpus study) and argued that the lack of coincidence was, deceptively, true in that it was observed only if a coarse grain is chosen (mere high or low preference at large). If, by contrast, frequency is allowed to show its role also in the tabulation of nonstructural parameters, like animacy or grammatical number, then the situation may turn out to be an entirely different one. Desmet et al. realised that the structure that had been put to the test in tens of studies was one involving two human nouns in the CNP (like our example (1), a classic one; Cuetos & Mitchell 1988). That was the inheritance of the original English studies, as in English the complicating that/who distinction could be avoided by choosing two animate NPs (with who as a relativiser). However, when Desmet et al. went to the corpora, they realised that the combination of two human NPs was statistically negligible (3%), which means that conclusions based on the processing of that pattern had been wrongly assumed to hold of the entire [CNP + RC] structure in most studies until then. When the animacy of the Ns was controlled, they reported an almost perfect match between the production and the reading data. Specifically, they uncovered a strong propensity to attach the RC high to the first NP when this coded an animate referent (animacy was not so strong in the second; see General Discussion). This led them to opt for a conception of the Tuning model that explicitly defends a fine grain of analysis: if such lexical properties of the nouns composing a syntactic structure have a role in shaping adjunction preferences, then syntactic processing is at least partially
driven by nonsyntactic forces. In their more recent study (Desmet et al., 2006), they confirmed the role of animacy and extended their research to another semantic property of nouns: the concrete/abstract distinction. In particular, they found out that when the first noun was both animate and concrete (daughter, boy, girl) as opposed to animate and abstract (government, staff, committee, and the like), attraction was strongest.

In this paper we report a piece of research conducted with a view to putting the Desmet et al.’s new findings about animacy to the test in Spanish. At present, the data on Spanish RC attachment counts is surprisingly sparse and there is little solid evidence on the extent to which Spanish corpus data and online preferences are aligned. First, we present data from a corpus and then two self-paced reading studies. Additionally, the comparison of the two studies with those on Dutch should also cast some cross-linguistic light on the role of this semantic property of noun phrases in the determination of attachment (i.e., syntactic) preferences.

CORPUS STUDY

The aim of the corpus study is to obtain information about the fine grain of Spanish [CNP + RCs], particularly on the animacy dimension. No other study of this kind has ever been done in Spanish before, as the pioneering work by Cuetos and collaborators utilised only a coarse-grain approach. Contrary to Desmet et al.’s most recent research (2002, 2006), we target only animacy, but we do so keeping the four possible combinations of animacy/inanimacy in mind: Animate NP1 + Animate NP2, Animate NP1 + Inanimate NP2, Inanimate NP2 + Animate NP1, and Inanimate NP1 + Inanimate NP2. We felt that since animacy was the isolable feature to focus upon, both in our corpus study and the online study, the contrast between the mixed conditions (Animate NP1 + Inanimate NP2, and Inanimate NP1 + Animate NP2) was essential. Also, since we are interested in establishing what type—if any—is actually affecting the coarse grain (i.e., what type is dominant), and whether this has any effect in ease of processing, we decided to study the remaining two types (the two animate NPs and the two inanimates). In previous work on another Iberian language, Galician (Fraga, Acuña, García-Orza, Teijido, & Franco-Grela, 2001; López, 2003), we obtained results pointing to a preponderance of the Inanimate NP1 + Inanimate NP2 condition which has not been verified in Spanish yet, and which is not exactly coincident with the Dutch data (where Inanimate NP1 + Animate NP2 was the dominant condition in the corpus).

In contrast with the four conditions, we will test in both production and reading, the most comparable online study in Dutch (Desmet et al., 2006)
kept all NP2s constant as Animate and focused on NP1 only, as shown in (2)–(5). Their English translations are:

The decisions of the president that … (2)
The documents of the president that … (3)
The organisations of the president that … (4)
The advisers of the president that (5)

METHOD

The corpus chosen contained 1.5 million words of both written and spoken Spanish. It can be accessed in the web page of the Real Academia Española (RAE; www.rae.es), the Spanish Royal Academy of Language. The RAE corpus offers unlimited possibilities on a variety of dimensions, ranging from size to register, all indexed by year, country, or topic. We chose 41% of written newspapers, 45% of magazines, and 14% of oral texts. The inclusion of a small sample of oral material had an exploratory purpose: We wanted to see if language which is meant to be spoken differs from written language in some hidden way that may impact attachment choices. We anticipate here that that was not the case and will not refer to this dimension of our work any further.

We looked for NP1 de/of NP2 que/that sentences, making sure that que was actually a relativiser, instead of a complementiser, which could adjoin to two immediately juxtaposed nominal hosts. Besides, we analysed the sample of sentences with a view to barring from the analysis any sentence that might be peculiar in any way. For instance, we discarded the sentence “nivel de acidosis (disminución del pH de la sangre) y de oxigenación que presenta el medio fetal en los momentos previos al parto” (level of acidosis (a decrease of blood pH) and oxygenation which is present in the foetal environment prior to delivery), because, among other things, the coordinating conjunction y/and makes it possible for the two nominals to attract the RC simultaneously. We also focused on preposition delof only, as we wanted to avoid dealing with the different predicative properties of the different prepositions (Frenck-Mestre & Pynte, 2000b; Gilboy et al., 1995).

In order to evaluate the role of animacy, we divided the corpus sentences in four different categories: A-A, I-A, A-I, and I-I (“A” meaning “animate” and “I” “inanimate”). We avoided instances where the animacy of the NPs might be problematic: e.g., nouns like government, which may denote animacy and abstractness at the same time. Decisions on the attachment choices (N1 or N2) were made by two referees (one of them a linguist). Examples which caused the referees to disagree were simply eliminated from the count. This happened very rarely, as the vast majority of choices were perfectly clear.
RESULTS

We obtained 777 sentences. The first thing to notice is that the category I-I is by far the dominating one in the corpus at large (559 instances, or 70% of the total). Next comes the category I-A (127 instances), then the category A-I (66), and finally the category A-A (with a mere 25 instances). Taken together, there is a significant preference for high adjunction, as 59% of all sentences were attached to the NP1 site, with 41% pointing to the NP2, $\chi^2(1) = 25.87, p < .001$. The results of the remaining tests confirm that this preference for the high adjunction occurs in all conditions, A-A, $\chi^2(1) = 4.84, p = .028$ A-I, $\chi^2(1) = 17.51, p < .001$, and I-I, $\chi^2(1) = 58.6, p < .001$, except in the I-A condition, where NP2 is preferentially chosen, $\chi^2(1) = 56.89, p < .001$, as a host for the RC (83.5%). This is in fact the most pronounced difference (see Table 1).

DISCUSSION

The corpus contains a moderate number of temporarily ambiguous sentences (777) with the structure NP1 del/of NP2 + RC whose adjunction is resolved towards the first NP in 59% of the cases and towards NP2 in 41% of them. This overall pattern of results confirms the tendency for Spanish to resolve this kind of ambiguity by looking up in the tree. The results obtained are similar to those reported in Cuetos et al. (1996, p. 176), with a corpus of 450,000 words, who found that the adjunction was resolved towards the first NP in 60% of the cases.

Regarding head types, preferences for NP1 structures are clear in three of the four conditions. Despite the general preference for NP1, animacy seems

TABLE 1

Number and percentages of NP1 and NP2 attachments in the corpus for each of the different animacy conditions

<table>
<thead>
<tr>
<th>Head type</th>
<th>NP1</th>
<th></th>
<th>NP2</th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>%</td>
<td>No.</td>
<td>%</td>
<td></td>
</tr>
<tr>
<td>A-A</td>
<td>18</td>
<td>72</td>
<td>7</td>
<td>28</td>
<td>25</td>
</tr>
<tr>
<td>A-I</td>
<td>50</td>
<td>75</td>
<td>16</td>
<td>25</td>
<td>66</td>
</tr>
<tr>
<td>I-A</td>
<td>21</td>
<td>16</td>
<td>106</td>
<td>84</td>
<td>127</td>
</tr>
<tr>
<td>I-I</td>
<td>370</td>
<td>66</td>
<td>189</td>
<td>34</td>
<td>559</td>
</tr>
<tr>
<td>Total</td>
<td>459</td>
<td>59</td>
<td>318</td>
<td>41</td>
<td>777</td>
</tr>
</tbody>
</table>

A-A: animate NP1, animate NP2; A-I: animate NP1, inanimate NP2; I-A: inanimate NP1, animate NP2; I-I: inanimate NP1, inanimate NP2.
to play a role in that only when the NP1 is inanimate and the NP2 animate is
the adjunction pattern reversed.

Over 70% of the sentences involved are of the category I-I, that is, the
category for which neither N is animate. Furthermore, the least frequent
condition is the A-A one, that is, precisely the condition used in the vast
majority of both online and offline studies to date. The classic sentence type
in (1), which revolutionised the syntactic ambiguities field, seems to be less
than frequent.

When our corpus study is compared to those carried out in Dutch,
similarities and differences arise. Three differences between the Dutch and
the Spanish data are worth keeping in mind. First, in the former the most
frequent pattern is that of condition I-A, whereas in the latter, it is the I-I
pattern that dominates the scene. Second, there is an overall preference for
NP1 in Spanish compared to the NP2 preference in Dutch. And third, in the
I-I condition, Spanish shows a robust NP1 bias, whereas Dutch opts for the
NP2 site.

Regarding similarities between these two languages, we can confirm that
animacy and overall attachment preferences both play a part in production
in Spanish as well as in Dutch. For instance, when the two NPs are animate,
the first NP attracts adjunction, both in Spanish and in Dutch. But, what is
more important, animacy seems to interact with the general attachment
preference, as can be seen from the fact that when the animate noun is NP2,
this also attracts the RC if the first NP is inanimate. This fact coincides with
data from Galician (Fraga et al., 2001; López, 2003) and Dutch (Desmet
et al., 2002, 2006) and becomes particularly interesting once we examine
adjunction preferences in the mixed conditions, A-I and I-A. In the first
place, high attachment drives the parse; in the second case there is an even
stronger preference for the lower site. In sum, only in this latter condition
does the parser behave in a way that runs counter to the general attachment
bias, disregarding the first site. The only possible explanation seems to be the
presence in the NP2 slot of an animate noun and the inanimateness of the
first noun.

EXPERIMENT 1

In Spanish no online experiment has ever been conducted in which the
animacy of the nouns involved was systematically examined. If the
comprehension data are aligned with the production data, we should find
shorter reaction times for NP1 disambiguated sentences in all cases except in
the I-A condition, where the opposite pattern of results should be obtained.
On the contrary, if there were characteristics of the Spanish language which
would render the animacy effects found in Dutch inoperative in this
language, then only the coarse-grain preference for the NP1 site should be manifested, even in the mixed category I-A.

Method

Participants. Eighty-three native Spanish speakers, who were enrolled in the courses of Logopedics in the University of Málaga, participated in the experiment for course credit. All participants had normal vision or corrected to normal and they were unaware of the aim of the study. Two participants, who had more than 10% of reading times as outliers, and nine, who failed in answering more than 40% of questions related to filler sentences correctly (see later), were eliminated. Seventy-two participants remained for analysis.

Materials and design. The experimental items comprised disambiguated sentences with the following structure: NP-V-NP1-de(of)-NP2-RC. Four types of NP1-NP2 relationships were considered: A-A, A-I, I-A, and I-I. Twelve sentences of each type were created making a total number of 48 sentences. Sentences in types I-A and A-I contained exactly the same words. Additionally, two versions of each sentence were created; in one case the relative clause was forced to a high attachment; in the other, to low attachment. Gender morphology was manipulated to create such sentences as shown in (6).

From these sentences four lists were built: Lists A and B contained 12 I-A (six disambiguated to NP1 and six disambiguated to NP2) and 12 A-A sentences (six disambiguated to NP1 and six disambiguated to NP2); Lists C and D contained 12 A-I (six disambiguated to NP1 and six disambiguated to NP2) and 12 I-I sentences (six disambiguated to NP1 and six disambiguated to NP2). Each list included also 50 filler sentences and four practice sentences.

Todos los amigos felicitaron / al cliente de la costurera / que sonreía emocionado / a la gente (6a’)

All the friends congratulated the client (masc.) of the tailor (fem.) who was smiling deeply moved (masc.) to the people (6a)

Todos los amigos felicitaron / al cliente de la costurera / que sonreía emocionada / a la gente (6b’)

All the friends congratulated the client (masc.) of the tailor (fem.) who was smiling deeply moved (fem.) to the people (6b)

A factorial design with two repeated factors was used. Attachment site, with two levels, and type of NP with four levels (A-A, A-I, I-A, I-I) were the variables.
Procedure. Participants were seated in front of a 14-inch monitor and a computer equipped with ERTS (Beringer, 1999). The experiment consisted of a self-paced reading task using a noncumulative moving window procedure. The task began with three screens giving instructions, which emphasised that participants had to read sentences which would appear as segments. Four sentences from the filler items were used as practice sentences. Experimental sentences were presented in five segments. The first included the whole sentence where letters were replaced by *; the second segment included the subject and the main verb of the sentence (NP-V); in the third the NPs were shown (NP1-NP2); the fourth showed the RC that contains the disambiguating word; and in the fifth the final part of the sentence was included to avoid wrap-up effects (usually a prepositional phrase). When the final segment of each filler sentence disappeared, a question followed by two alternatives was shown on the screen. Participants had to press one of the two shift lateral keys to select their answer, the key corresponding to the side of the chosen answer. Using this procedure, we encouraged participants to read for comprehension.

Half of the participants were exposed to Lists A and C, and the other half to Lists B and D (see Materials section). In this way participants were exposed only to one version (high or low attachment) of each experimental sentence and read sentences with the four types of NP relationships. Within each list sentences were presented in a different random order for each participant. Between each list a pause of a few minutes was introduced for resting. The entire experiment took about 25 min.

Results

Table 2 presents the mean reading times in the critical segment (where disambiguation takes place) in the different conditions. An interaction between sentence type and attachment should be found (with faster reading times in NP1 sentences in three out of the four sentences types and faster reading times in NP2 in the other (I-A) condition) in the disambiguation region that parallels the preferences found in the corpus data. Nonetheless, to further analyse the relationship between corpus and reading times in the following ANOVAs, we treated the head conditions, (i.e., the NP type variable) as having four levels of corpus bias. That is, NP type was treated as a quantitative variable, as NP type conditions were substituted by the NP1 bias of each head type in the corpus (a bias of 72% replaced the A-A condition, one of 75% replaced the A-I condition, one of 16% replaced the I-A condition, and one of 66% replaced the I-I condition). In this way, the interaction between attachment site and the linear component of the NP type variable tests whether increasing corpus bias changes the online
attachment latency preference. This analysis is the most sensitive index of the predicted relationship between the corpus and the reading time data.\(^1\) Put differently: Because the present experiment was concerned with the numerical correspondence between corpus values and attachment preferences, the interaction between attachment site and the linear component of NP type (i.e., the linear polynomial contrast) is more informative than the usual interaction yielded by the ANOVA.

A 2 (attachment site: NP1 vs. NP2) × 4 (type of NP: A-A: 72%; A-I: 75%; I-A: 16%; I-I 66%) analysis of variance with participants (1) and items (2) as random factors was conducted. We found a significant effect of attachment site, \(F_1(1, 71) = 4.41, p = .039; F_2(1, 44) = 6.21, p = .016,\) indicating that participants were faster in NP1 than in NP2 disambiguated sentences. A significant effect of NP type was also found, \(F_1(3, 213) = 13.25, p < .001; F_2(3, 44) = 3.67, p = .02.\) Planned comparisons revealed that I-I sentences were read faster than the remaining types of sentences in the participants analyses (all \(p_1 s < .01\), but only faster than the A-I sentences in the items analyses, \(p_2 = .017.\) However, this effect is not particularly relevant as sentence segments differed in length between the different types of sentences (however, recall that segments in the A-I and I-A sentences have exactly the same words, see later). Numerically, the reading time differences between NP1 and NP2 sentences were similar to the data found in the corpus. More importantly, the analysis of the Attachment site × NP type-linear interaction revealed significant differences by items, \(F_2(1, 46) = 8.97, p < .01,\) that were marginally significant by participants, \(F_1(1, 71) = 3.2, p = .07.\) This interaction suggests an alignment between attachment preferences while reading and preferences in the corpus when animacy is taken into account.

Planned comparisons considering attachment preferences were carried out for each NP type. Results showed marginally significant differences in the A-A (\(p_1 = .05\) and \(p_2 = .1\)) and in the I-I sentences (\(p_1 = .07, p_2 = .1,\) not approaching significance in the A-I and I-A conditions.

| TABLE 2 |
|---|---|---|---|---|
| | A-A | A-I | I-A | I-I |
| **M** | **SD** | **M** | **SD** | **M** | **SD** | **M** | **SD** |
| NP1 | 1274.41 | 238.84 | 1345.76 | 287.30 | 1314.21 | 273.09 | 1174.01 | 246.76 |
| NP2 | 1328.05 | 315.02 | 1365.54 | 283.95 | 1309.44 | 279.60 | 1212.37 | 253.47 |

\(^1\) We are indebted to Don Mitchell for suggesting this type of analysis.
A comparison between I-A and A-I sentences was carried out to directly test the role of animacy. This analysis is particularly relevant, as these sentences were composed exactly of the same words, but in different order. Moreover, the corpus analysis showed contrasting preferences for this type of sentences: more NP1 sentences for the A-I condition and an NP2 advantage for I-A sentences. Hence, we ran separate analysis for these sentences taking into account attachment site (NP1 vs. NP2) and NP type (A-I vs. I-A) as variables. The ANOVA did not show main effects of NP type, $F_1(1, 71) = 2.17, p = .14; F_2(1, 11) = 1.68, p = .22$, nor attachment site (both $Fs < 1$). The interaction was significant by items, $F_2(1, 11) = 7.78, p = .018$, but not by participants, $F_1(1, 71) = 0.58, p = .45$. Data showed that when I-A and A-I conditions were analysed separately, no clear evidence of the role of animacy was found.

**Plausibility study.** There is a factor that could be biasing the attachment preferences found in our online study: the plausibility of NP1 and NP2 as subjects of the RC. It has been shown (e.g., Bock, Loebell, & Morey, 1992) that animate noun phrases are more plausible subjects than inanimate ones, so a plausibility rating study was run to check the influence of this factor in our results. Experimental sentences were rephrased into active sentences where the NP1 and NP2 were the subjects of the verbs in the RC, now construed as the main verb in the sentence. That is, for the example in (6a) the corresponding sentence was “El cliente de la costurera sonreía emocionado a la gente” (The client (masc.) of the dressmaker (fem.) was smiling deeply moved (masc.) to the people) and for (6b) “la costurera sonreía emocionada a la gente” (The dressmaker (fem.) was smiling deeply moved (fem.) to the people). Thirty-six participants took part in this study. They rated the sentences on a 5-point scale (1 = “implausible”, 5 = “plausible”). According to a Latin-square design, four lists were built with the experimental sentences. Participants were exposed only to one version of each sentence. ANOVAs with type of NPs (A-A, I-A, A-I, I-I) and NP attachment sentences (those coming from NP1 vs. NP2 experimental sentences) were run over participants and items. Results showed that NP2 sentences were rated as more plausible than NP1 sentences and this difference was significant by participants, $F_1(1, 35) = 4.86, p = .03; F_2(1, 44) = 2.98, p = .09$. No effects of NP type, $F_1(3, 105) = 2.04, p = .11; F_2(1, 44) = 0.82, p = .45$, nor interactions, $F_1(3, 105) = 1.26, p = .29; F_2(3, 44) = 0.46, p = .71$ were found.

Planned comparisons showed a higher plausibility of the NP2 compared to the NP1 sentences both in the A-A ($p_1 = .04; p_2 = .05$) and I-A sentences ($p_1 = .02; p_2 = .09$). Mean plausibility ratings for the A-A sentences were for NP1 = 4.67 and for NP2 = 4.83. In the I-A condition these values were 4.72 and 4.87, respectively. No differences arose in the I-I (NP1 = 4.65;
NP2 = 4.66) and A-I (NP1 = 4.64; NP2 = 4.70) conditions. When the plausibility ratings of I-A and A-I sentences were analysed taking into account these types of NPs alone, results showed main effects of attachment site (NP2 sentences rated as more plausible) in the analysis over participants, $F_1(1, 45) = 4.87, p = .03; F_2(1, 11) = 3.01, p = .11$. Effects of head type were marginal, $F_1(1, 35) = 3.58, p = .06; F_2(1, 11) = 2.77, p = .12$, and no interaction effects were found, $F_1(1, 35) = 1.29, p = .26; F_2(1, 11) = 0.23, p = .63$.

The results of this plausibility rating study suggest that the overall preference for NP1 sentences in the reading experiment is not caused by plausibility confounds. Despite the higher plausibility of NP2 sentences, participants attached the RC to NP1. Moreover, we cannot rule out the possibility that the higher plausibility of the NP2 heads could be diminishing the overall NP1 preferences in our reading experiment, making them less strong than they actually are (remember that although a main effect of attachment site towards NP1 was found, no differences arose in planned comparisons when this variable was in the four types of NP sentences). In other words, plausibility could be playing a modulating role that would avoid the finding of clearer NP1 attachment preferences in our data.

Discussion

Overall, the Attachment site × NP type-linear interaction, significant by items and marginally significant by participants, suggests that the online preferences were not homogeneous and were closely related to those found in the corpus. This interaction showed that attachment preferences while reading correspond with the frequency of the high and low attachment structures found in our corpus. There was a statistically significant main effect of attachment site, showing that globally NP1 sentences were read 27 ms faster than NP2 sentences. However, when attachment preferences regarding head type were considered, despite the interaction reported, none of the numerical preferences (for NP1 sentences in all the cases except the I-A condition, in which a slight numerical preference for NP2 was found) were statistically significant. Moreover, when the Head type × Attachment site interaction was considered in sentences that only differ in the order of head types (see sentences A-I vs. I-A), the interaction was only significant over items. These data make the evidence in favour of the correspondence between corpus and reading data less strong. One factor that could be responsible for this unclear picture is plausibility, as the higher plausibility of NP2 sentences could be moderating the numerical NP1 preferences found in three out of the four conditions.

In sum, our data in this experiment suggest that animacy could be playing a relevant role in Spanish online disambiguation of relative clauses, as
indicated by the Attachment site x NP type-linear interaction. However, this conclusion should be taken with care, as there are no attachment preferences in any type of sentence and no differences are found between I-A and A-I sentences regarding attachment.

EXPERIMENT 2

With the aim of obtaining clearer results, a second experiment was carried out replicating the first experiment but with two main differences. First, the segment by segment self-paced reading task was substituted by a word by word self-paced procedure with the aim of having a more sensitive method to detect attachment preferences. This presentation also minimises the undesired effects that the previous segmentation could have had on attachment preferences. Note that according to Gilboy and Sopena (1996) the segmentation used in Experiment 1 could be favouring NP1 attachment (but see Carreiras & Clifton, 1999). Second, the items were matched for plausibility, so any reading time differences observed in this second experiment cannot be due to differences in the plausibility of NP1 and NP2 as subjects of the RC.

Method

Participants. Seventy-four native Spanish speakers, who were enrolled in the courses of Psychology at the University of Málaga, completed the experiment. All participants had normal vision or corrected to normal, were unaware of the aim of the study, and participated as volunteers. Two participants who failed to correctly answer more than 40% of the questions related to filler sentences were eliminated. Seventy-two participants remained for analyses.

Materials and design. We used the same types of sentences as in Experiment 1, but they were slightly modified till we obtained similar ratings of plausibility between the NP1 and the NP2 sentences in each of the head type conditions. A similar procedure to that followed in the plausibility ratings of Experiment 1 was employed in the plausibility ratings of the sentences to be employed in Experiment 2. Forty-eight sentences were rephrased into active sentences where the NP1 and NP2 were the subjects of the RC. According to a Latin-square design, seventy-two participants (different from those that took part in the experiment) rated only one version of each sentence on a 5-point scale. ANOVAs showed no differences between NP1 and NP2 sentences (mean plausibility ratings respectively, 4.90 and 4.89), $F_1(1, 71) = 0.85, p = .35$; $F_2(1, 44) = 0.77, p = .38$. Although I-I sentences had lower plausibility ratings than the other head type sentences,
there were no effects of head type, $F_1(3, 71) = 2.4, p = .06$; $F_2(3, 44) = 1.56, p = .21$, nor interactions, $F_1(3, 213) = 1.3, p = .27$; $F_2(3, 44) = 0.08, p = .49$. Planned comparisons did not show differences in plausibility between NP1 and NP2 as heads of the RC in any of the sentence types. When only I-A and A-I sentences were considered, the analyses did not show main effects of head type, NP attachment nor interactions.

Procedure. The procedure was similar to that in Experiment 1, with a critical difference: In Experiment 1 sentences were presented phrase by phrase in five segments, whereas presentation in Experiment 2 was word by word. This presentation allowed us to measure the reading time in the disambiguating word alone and not as in Experiment 1, where that word was part of a segment that included the personal pronoun and the verb of the RC. As in Experiment 1, participants were exposed only to one version (high or low attachment) of each experimental sentence and read sentences with the four types of NP relationships. The experiment was administered in two sessions separated by 1 week. Each session lasted about 15 min.

Results

For purposes of analysis, items were divided into six regions, as follows: The first region includes the subject and the main verb of the sentence; the second region includes the CNP; the third includes the beginning of the RC, usually the pronoun and the verb of the RC sentence; the fourth region includes the disambiguating word; the fifth includes the two words following the disambiguating word, and the sixth the rest of the sentence. An example is presented in (7).

Todos los amigos felicitaron / al cliente de la costurera / que sonreía / emocionado / a la gente (7)

All the friends congratulated / the client (masc.) of the dressmaker (fem.) / who was smiling / deeply moved (masc.) / to the / people (7)

Table 3 presents the mean reading times from the beginning of the RC sentence crossing head type and attachment site.

Although our main interest was in the disambiguating word, analyses were run from the beginning of the RC (Region 3) to the end of the sentence (Region 6) with participants and items as random factors. As signalled in Experiment 1, NP type was considered as a quantitative variable and the Attachment site × NP type-linear interaction as the test of the correspondence between corpus bias and attachment preferences while reading.

Analyses in the third region, the beginning of the RC, showed that although reading times in NP2 were faster than in NP1 sentences, this
difference did not approach significance, $F_1(1, 71) = 2.6, p = .11; F_2(1, 44) = 1.4, p = .23$. Type of NP was not significant (all $F_1$ and $F_2 < 1$). The interaction between the linear component of NP type and attachment site was not significant ($F_1$ and $F_2 < 1$). As expected, planned comparisons in each NP type sentence showed no differences between NP1 and NP2 sentences in this region (all $p_s > .25$).

The fourth region, the disambiguating word, showed no effects of attachment site ($F_1$ and $F_2 < 1$) and marginally significant effects of NP type, $F_1(3, 213) = 2.8, p = .07; F_2(3, 44) = 2.8, p = .05$. The Attachment site $\times$ NP type-linear interaction was not significant ($F_1$ and $F_2 < 1$). Planned comparisons revealed that participants were faster reading the disambiguating word in NP1 sentences when the CNP consisted of an animate NP1 and an inanimate NP2, that is, in the A-I condition ($p_1 < .01; p_2 = .06$). In the remaining animacy conditions no differences between NP1 and NP2 sentences arose (all $p_s > .5$).

In the fifth region, the one that includes the two words following the disambiguating word, main effects of attachment site were found, $F_1(1, 71) = 16.9, p < .001; F_2(1, 44) = 3.6, p = .06$: NP1 sentences were read about 30 ms faster than NP2 sentences. Effects of sentence type were also found in the analysis over participants, $F_1(3, 213) = 5.2, p < .01; F_2(3, 44) = 0.5, p = .7$. Finally, and more importantly, the interaction between the linear component of NP type and attachment site was significant by participants, $F_1(1, 71) = 6.8, p = .01; F_2(1, 46) = 2.2, p = .14$, showing a linear change in attachment preferences related to corpus bias in the different animacy conditions. Planned comparisons showed differences between NP1 and NP2 that were only significant over participants, in the A-A sentences ($p_1 = .02; p_2 = .18$), in the I-I sentences ($p_1 = .04; p_2 = .39$), and in the A-I sentences ($p_1 = .01; p_2 = .07$). In the I-A condition, where the NP2 sentences were read 25 ms faster than the NP1 sentences, this difference was not significant ($p_1 = .2; p_2 = .81$).

<table>
<thead>
<tr>
<th>Sentence type</th>
<th>Attachment site</th>
<th>Region 3</th>
<th>Region 4</th>
<th>Region 5</th>
<th>Region 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-I NP1</td>
<td>890</td>
<td>431</td>
<td>869</td>
<td>675</td>
<td></td>
</tr>
<tr>
<td>A-I NP2</td>
<td>878</td>
<td>462</td>
<td>939</td>
<td>695</td>
<td></td>
</tr>
<tr>
<td>I-I NP1</td>
<td>882</td>
<td>424</td>
<td>811</td>
<td>673</td>
<td></td>
</tr>
<tr>
<td>I-I NP2</td>
<td>862</td>
<td>418</td>
<td>839</td>
<td>724</td>
<td></td>
</tr>
<tr>
<td>I-A NP1</td>
<td>857</td>
<td>411</td>
<td>870</td>
<td>635</td>
<td></td>
</tr>
<tr>
<td>I-A NP2</td>
<td>848</td>
<td>410</td>
<td>845</td>
<td>638</td>
<td></td>
</tr>
<tr>
<td>A-A NP1</td>
<td>884</td>
<td>415</td>
<td>823</td>
<td>704</td>
<td></td>
</tr>
<tr>
<td>A-A NP2</td>
<td>875</td>
<td>407</td>
<td>864</td>
<td>713</td>
<td></td>
</tr>
</tbody>
</table>

Mean reading times in Experiment 2 for the RC particle and verb (Region 3), the disambiguation word (Region 4), the following two words (Region 5), and the rest of the sentence (Region 6) as a function of animacy and attachment site.
The differences described in the fifth region did not continue in the sixth region.\textsuperscript{2} Main effects of NP type, $F_1(3, 213) = 6.7$, $p < .01$; $F_2(3, 40) = 0.3$, $p = .8$, were found. Attachment was only marginally significant by participants, $F_1(1, 71) = 3.8$, $p = .06$; $F_2(1, 40) = 2.9$, $p = .1$, and the interaction including the linear component of NP type did not show any effect ($F_1$ and $F_2 < 1$). Planned comparisons in each of the NP type sentences showed no differences between NP1 and NP2 sentences in three out of the four conditions (all $ps > .5$). However, in the I-I condition there was a significant preference for NP1 sentences ($p_1 = .05$; $p_2 = .04$).

As in Experiment 1, separate two-way ANOVAs were run for I-A and A-I sentences. In the third region, as expected, there were no effects of attachment site nor interactions between attachment site and NP type (all $F$s < 1.2). Data showed that I-A sentences were read about 30 ms faster than A-I sentences. This difference was only significant in the analysis over items, $F_1(1, 71) = 1.6$, $p = .21$; $F_2(1, 11) = 9.9$, $p < .01$. In the fourth region a marginally significant effect of attachment site, favouring NP1, was found in the analysis over participants, $F_1(1, 71) = 3.8$, $p = .05$; $F_2(1, 11) = 1.2$, $p = .3$. Again I-A sentences were read faster than A-I sentences, $F_1(1, 71) = 4.4$, $p = .04$; $F_2(1, 11) = 7.6$, $p = .02$. More importantly, the interaction was significant in the analysis over participants, $F_1(1, 71) = 5.1$, $p = .02$; $F_2(1, 11) = 1.28$, $p = .28$. Planned comparisons showed that the NP1 preferences were significant in A-I sentences ($p_1 < .01$; $p_2 = .2$) and were absent in the I-A sentences (both $ps > .9$). In the fifth region, there were significant effects over participants of NP type, $F_1(1, 71) = 4$, $p = .04$; $F_2(1, 11) = 3.49$, $p = .08$ (I-A sentences were read 46ms faster than A-I sentences) and attachment site, $F_1(1, 71) = 4.8$, $p = .03$; $F_2(1, 11) = 0.93$, $p = .35$. The interaction between these variables was significant in the analysis over participants, $F_1(1, 71) = 4.6$, $p = .03$; $F_2(1, 11) = 2.14$, $p = .17$. Again, the NP1 preferences were significant in A-I sentences ($p_1 < .01$; $p_2 = .2$) and were absent in the I-A sentences (both $ps > .2$). Finally, in the sixth region there were effects of NP type, $F_1(1, 71) = 5.2$, $p = .02$; $F_2(1, 10) = 24.7$, $p < .001$, but no effects of attachment site (both $Fs < 1$) nor interactions (both $Fs < 1$).

\textbf{Discussion}

The most noteworthy aspects of the data start in the disambiguating region (Region 4). Although there were no effects of attachment site nor an interaction between this factor and the linear component of NP type, the planned comparison showed a preference for NP1 in the A-I condition. In

\textsuperscript{2} Due to differences in length between sentences, two I-I, one A-I, and one I-A sentences only have five regions; hence, they were not included in these analyses.
that specific configuration, the NP1 bias that has been generally reported for Spanish shows at the earliest possible time. It is also as well to note that the comparison between the two symmetrical conditions (A-I and I-A) revealed an interaction between attachment site and NP type (nonsignificant by items). There were preferences for NP1 in the A-I condition and no commitment for any structure in the I-A condition. The interaction showed the role of animacy in pushing attachment preferences. However, this factor was not enough to obtain NP2 preferences in the A-I condition.

The role of animacy and attachment preferences were more clearly shown in Region 5, where NP1 and NP2 sentences had exactly the same words. However, the finding of differences regarding attachment site in this region is not surprising. It is a well-known fact about word by word self-paced reading tasks that the processing of each word can spill over into the time in which subsequent words are presented, because participants can fall into a routine of tapping the buttons at regular intervals (e.g., Just, Carpenter, & Wooley, 1982; Mitchell, 1984). In this region, there were main effects of attachment site (with faster NP1 sentences) and sentence type. The Attachment × NP type-linear interaction was significant in the analysis over participants, showing the correspondence between the preferences found in the corpus and those in the online tasks. Revealingly, there were significant preferences for NP1 in all conditions (but not in the items analysis) except in the I-A condition, where the NP2 site is the preferred option (some 25 ms faster than NP1), despite not reaching significance. The comparison between the two mixed animacy conditions showed interaction effects as well, but again, only in the analysis over participants.

In Region 6 the interaction effects disappear and only a significant preference for NP1 is still observed in the I-I condition.

The comparison between conditions A-I and I-A casts an interesting effect: starting in Region 4, the I-A condition is consistently read faster than the A-I one. Although one cannot rule out an explanation for this finding based on the intrinsic cognitive advantage of such a configuration (it might be easier to locate objects by reference to animate beings than to locate animate referents via the position that objects take relative to them), it is surely of interest that the faster condition is the one that appears more often in the corpus (almost twice as much).

Overall, these analyses point out the relevance of the type of NP in RC attachment. When the animate head is in NP1 position, preferences to attach the RC to this head were strong. On the contrary, the presence of an animate head in the NP2 position, together with an inanimate head in the NP1, made the usual preference for NP1 disappear, and although significant NP2 preferences were not observed, there was a clear advantage towards that attachment site.
All in all, it would appear that the word-by-word presentation methodology has accentuated the preferences found in our first experiment, although such preferences were more conspicuously observed in the two words following the disambiguating region now. However, the lack of significant differences in the analyses over items in our data might undermine the generality of our conclusions vis-à-vis other materials. In this sense, the small number of sentences in each condition might have played some kind of a role in the overall pattern of results. In any case, the results in this experiment do provide modest support for a fine-grain version of Tuning that views animacy as a determinant of online attachment preferences.

GENERAL DISCUSSION

A previous study of Spanish relative clause ambiguities showed correspondences between corpus data and attachment preferences while reading, supporting a coarse-grain version of the Tuning model (Mitchell et al., 1992). However, later studies on Dutch have clearly revealed that such a grain of analysis cannot explain attachment preferences in that language (Mitchell & Brysbaert, 1998). More recently, new studies have shown that the correspondence between reading times and corpus data does actually take place, but only when the animacy of the NPs is taken into account. The new data have been used to defend a fine-grain version of Tuning according to which not only the relative frequency of syntactic structures but also the relative frequency and semantic make-up of the lexical hosts of the relative clause determine attachment biases (Desmet et al., 2006). The first aim of this research has been to test whether a similar fine-grain hypothesis should be defended for Spanish. A corpus analysis and two self-paced reading experiments were carried out manipulating the animacy of the NPs. Results show preferences for NP1 sentences in three out of four conditions and NP2 preferences (although nonsignificant) for the inanimate-animate condition (that is, with animacy on the second site). The interaction between attachment and the linear component of animacy (this variable representing the percentage of NP1 preferences found in each NP type condition in the corpus) has shown a correspondence between corpus data and reading times data that is in line with a fine-grain version of the Tuning hypothesis. In essence, then, the first general conclusion that one can draw from the comparison between the corpus data and the reading data presented here is that they are basically consistent with the studies on Dutch. Specifically, Desmet et al. (2002) showed in a completion study that when NP1 is animate in that language it carries the adjunction of the RC. And Desmet et al. (2006) confirmed the same results in comprehension (eyetracking) and proved that whenever animacy was coupled with concreteness on the first
NP, high adjunction was even more pronounced. They obtained interactions by subjects and items, and we obtained an interaction by items (marginal by participants) in our first experiment and another interaction by participants in our second experiment. The analyses in both studies show that overall there is something going on with animacy in both languages.

In fact, the role of animacy in shaping grammatical form is indeed well-known in linguistics. Animacy runs through most agreement systems based on morphological features (Corbett, 2006), agreement systems being one of the central pillars of the organisation of the clause. Animacy is also linked to the grammar of topicality (topics are mostly animate and definite; Lambrecht, 1994; Goldberg, 2006) and various other formal properties of sentences have been shown to correlate with animacy hierarchies. Take, for instance, the Saxon Genitive. Unlike German, English genitive phrases are sensitive to semantic variables like animacy in the sense that the first noun in the noun phrase must exceed or at least be equal to the second noun in terms of animacy (Pete’s bike vs. *the bike of Pete; cf. Hawkins, 1981). Desmet et al. themselves remind us that semantic properties of nouns in general, and the relationships between different nouns at the semantic level, are important in syntactic processing as well. For instance, Gibson, Pearlmutter, and Torrens (1999) showed that the very same noun (planet) could attract RCs more or less depending on the other words it cooccurred with, even while keeping its own position constant. In production, there is also a strong tendency to code animate entities in subject positions (Bock et al., 1992; Itagaki & Prideaux, 1985; Piñeiro, Fraga, García-Orza, & Acuña, 2007). Desmet et al. (2006) argue that their own findings are better aligned with models that integrate frequency information at different levels of organisation, like Jurafsky’s (1996) probabilistic model based on conditional probabilities, Tabor, Juliano, and Tanenhaus’ (1997) dynamical system, or McRae, Ferretti, and Amyote’s (1997) thematically driven model. This in itself is a clear defence of the finest grain of analysis within the ample exposure-based postulates that characterise the Tuning framework (e.g., Mitchell et al., 1995).

An interesting aspect of the accumulated evidence provided by the Dutch and now the Spanish fine-grain corpus studies is that they confirm the important fact that indeed the animate-animate dimension is, by far, the least frequent in both languages (3% in Dutch, 3.2% in Spanish). More research is needed to see what happens in other languages, but the fact that Spanish and Dutch are not extremely similar overall and yet still show the same clear tendency suggests that we might be witnessing a solid cross-linguistic trend. Since almost all conclusions reached in the last 20 years or so about [CNP + RCs] have been based on the A-A pattern, one must declare them not too representative. This is an important finding, first uncovered by
the team working on Dutch and now firmly confirmed by the new Spanish data.

There are also differences between the Dutch and the Spanish data, and we should bear them in mind. The Inanimate-Abstract NP1 + Animate NP2 should in all logic show the most robust of the NP2 adjunction preferences. However, Desmet et al. (2006) could only show a nonstatistical trend in that direction in their eyetracking data. This is all the more surprising given the fact that the I-A condition is overall also the predominant pattern in the corpus analysis in Dutch (not in Spanish: I-I). In the second place, we still need to explain why the animate-animate type pulls adjunction to the first NP (despite overall corpus preference for NP2). The match between the corpus and the reading data in that category only pushes explanation one step back, as we still need to know why, when the two Ns are kept animate, we should still find an advantage for the higher site, one that is counter to the general tendency of the language. If animacy/concreteness is the key, one would expect no strong biases there, and given the overall NP2 preference of Dutch, maybe even a slight NP2 bias. The fact that animacy seems to work for NP1 only is therefore somewhat hard to explain. In that respect, the preference for NP1 in Spanish Animate + Animate phrases could indeed be attributed to coarse-grain level frequency effects, as indeed, overall, NP1 is preferred in Spanish (59%). It is interesting that when animacy is taken out of the equation (in the I-I condition), the Spanish corpus data show a sturdy NP1 preference. At that level of analysis, it is the Dutch data that remain mysterious, therefore. In Spanish, the semantic make-up of the lexical heads may explain inclinations of the RC to look to either noun host when this is animate. It is true that these inclinations do not solidify into something statistically significant when animacy falls on the second site, but the mere fact that the otherwise strong NP1 preference is neutralised precisely in the condition where there is an animate second noun is too obvious to ignore. And, despite not reaching significance, those NP2 sentences are actually read faster in the I-A condition than in the A-I condition. In Dutch, however, those animacy effects are very strong, but only for NP1s. Desmet et al. make reference to the fact that the object position of all the experimental sentences might have something to do with the NP1 bias. However, that is the standard function used across experiments in the last 20 years, so cross-linguistic differences cannot depend solely on that. And, in any case, one should still be able to specify what makes the object function behave that way.

One reason why the Dutch data might be more difficult to interpret as far as animacy on the second site is concerned might have to do with the NP1 singular + NP2 plural disambiguation that characterised half of the trials (the other half being the reverse combination of morphological number: plural + singular). In production at least, the phenomenon of attraction is a
solid one (Bock & Eberhard, 1993; Bock, Eberhard, & Cutting, 2004; Bock, Nicol, & Cutting, 1999; Eberhard, 1997; Franck et al., 2004; Franck, Lassi, Frauenfelder, & Rizzi, 2006; Franck, Vigliocco, & Nicol, 2002). It occurs when a plural local noun stands in the way of an agreement relationship between a verb and a more distant singular one, and produces frequent errors like *the key to the cabinets are in the drawer. Attraction is particularly strong when the plural noun follows a singular head, not vice versa, and has been attributed to the cognitive salience and markedness of the plural feature relative to the unmarked nature of the singular (Harley & Ritter, 2002). In fact, in a recent series of experiments on attraction, Thornton and MacDonald (2003) studied the role of plausibility in subject-verb agreement operations in English and claim to have found robust effects not only in production but also in comprehension. The plausibility of the verb was manipulated so that either the two nouns of a complex subject NP could be plausible passive subjects (the album by the classical composers ... BE praised) or only the head noun could be so (the album by the classical composers ... BE played). There were more agreement errors in production and longer RTs at the verb in comprehension when both nouns were plausible subjects than when only the head noun was plausible. In all cases the second noun of the complex NP was plural. It is difficult to predict if such a peculiarity of the Dutch materials may have interfered with attachment choices or not, but the strength of attraction is certainly considerable (especially in production), so interference should not be discarded. In our experiments both nouns were always singular—the unmarked, default scenario.

In sum, independently of the relationship between our studies and those on Dutch, our work has sought to cast light on the effects of fine-grain dimensions on the final determination of adjunction (i.e., syntactic) operations in Spanish. After more than 20 years of research on this kind of ambiguity, there are still significant gaps in our knowledge of the processing of this structure. It is clear that corpus studies such as that of Mitchell et al. (1992), which targeted only a coarse grain of analysis (overall preference for NP1 or NP2) using a sample of only half a million words, are no longer enough, so our own corpus study tripled the size of that pioneering work and focused on the precise lexical compositions of the nouns involved. With only one possible nonsyntactic variable controlled for (animacy), our corpus analysis cast a series of important facts: In the first place, there is a very pronounced asymmetry of NP types, with the I-I type taking up 70% of the CNPs overall; in the second place, there is a clear NP1 tendency in all the configurations studied except, revealingly, in the I-A type; in the third place, overall, in the absence of animacy (the I-I condition), Spanish turns to NP1 (remember that Dutch opts for NP2); finally, our corpus aligned itself with the Dutch data to confirm that the A-A dimension
is not really representative. Additionally, our online experiments provide some interesting lines of convergence with the corpus data. The fact that the I-A configuration is read faster than its mirror image, the A-I pattern, is surely of interest, especially since that entails a precise match with the corpus counts. The most interesting finding is that the general NP1 preference vanishes precisely when an animate second referent does not reside in the NP1. The numerical trend towards NP2 in the I-A condition is crucial in that it shows that the adjunction preferences which are solid across the various experiments on Spanish carried out to date (Carreiras, 1992; Carreiras & Clifton, 1993, 1999; Cuetos et al., 1996) can actually be modulated by semantic variables first uncovered by the Desmet et al. team for Dutch and now confirmed in our Spanish data (see also Soares, Fraga, Comesaña, Piñeiro, & Pinheiro, 2008, for recent data supporting the role of animacy in reading and completion studies in European Portuguese). It is also surely significant to find out that when animacy falls on the first site (which is already the preferred one independently of animacy), we have found the earliest reaction time response in our second experiment, the one with a careful control of plausibility. Overall, it seems that a meticulous, fine-grained examination of a series of nonstructural parameters like animacy (or concreteness or topicality) is likely to reveal interesting facets of the processing of adjunction ties online. Scrutinising corpora in search of such fine grains may reveal sturdy trends, and these often force one to ask questions that act as sparks for future research.

Original manuscript received December 2006
Manuscript accepted October 2008
First published online March 2009

REFERENCES


