Relative Clause Attachment in English and Spanish: A Production Study

Anouschka Bergmann*, Meghan Armstrong** & Kristine Maday*

Linguistics Department*, Department of Spanish and Portuguese**
Ohio State University
{anouschka; meghan}@ling.osu.edu; maday.3@osu.edu

Abstract

The Implicit Prosody Hypothesis assumes that differences in individual languages' attachment preferences for syntactically ambiguous sentences are due to the languages' different prosodic systems. For example, when a relative clause modifies a complex noun phrase (NP1 NP2 RC), a prosodic break after NP1 is said to trigger a low attachment interpretation, while a break after NP2 triggers high attachment. English and Spanish participants read aloud sentences of this type, but showed no correlation between prosodic phrasing pattern and attachment choice. Both English and Spanish readers pronounced the majority of sentences with the strongest prosodic break following NP2. However, responses to comprehension questions immediately following each production showed that Spanish speakers preferred a high attachment interpretation, while English speakers preferred low attachment. Our findings provide evidence against a prosodic account of overall attachment preferences for this construction and provide insights into the mechanisms of reading aloud.

1. Introduction

Late Closure, the idea that lexical items are attached into the clause or phrase currently being processed, if possible [1], is one of the most durable and influential assumptions in models of human sentence processing. According to Late Closure, a constituent with two possible attachment sites should preferably attach locally or low in the syntactic tree. The sentence *Someone shot the servant of the actress who was on the balcony*, for example, is globally ambiguous and the relative clause (RC) *who was on the balcony* can either attach low (i.e. the actress is on the balcony) or high (i.e. the servant is on the balcony), as illustrated in (1).

(1)

…the servant of the actress who was…
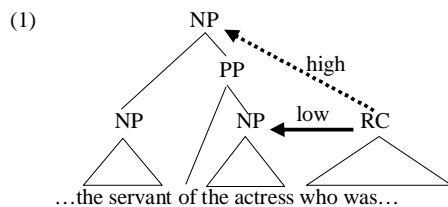
Late Closure is assumed to allow for more rapid and efficient attachment than Early Closure because fewer effortful shifts between levels of the syntactic tree are necessary [2]. Since this reasoning is entirely general, Late Closure is assumed to be fully innate and universal (cf. [3]). And, in fact, the tendency to attach low is well documented for a wide variety of constructions and across languages (e.g. [1]).

However, the syntactic construction shown in (1), where an RC modifies a complex noun phrase (NP), does not show a universal Late Closure preference. In 1988 Cuetos and Mitchell showed an Early Closure preference for this construction in Spanish [4], which has since then been confirmed in numerous studies (e.g. [5], [6]). Furthermore, an Early Closure preference for this structure has been found in many other languages, including Afrikaans, Croatian, Dutch, French, German, Greek, Japanese, Polish, Brazilian Portuguese and Russian (cf. [7]). These results pose a direct challenge to a universal Late Closure parsing principle.

Many attempts to save Late Closure have been proposed since. One of the most popular is Fodor's Implicit Prosody Hypothesis (IPH; e.g. [3], [8]). The IPH assumes that Late Closure is universal, but that differences in languages' prosodies can cause an Early Closure preference. Consistent with studies of silent reading, the IPH assumes the projection of a prosodic contour (implicit prosody) onto written text. Furthermore, implicit prosody and overt prosody, the actual production of a sentence, are assumed to be the same for a given speaker and sentence.

IPH studies assume that a prosodic boundary after NP1 is consistent with a Late Closure interpretation (e.g. [the servant] [of the actress who was…]) because NP2 (*the actress*) is prosodically grouped with the RC. However, a boundary after NP2 is consistent with an Early Closure interpretation (e.g. [the servant of the actress] [who was…]) because NP2 is separated from the RC.

This suggests that speakers of low attaching languages frequently produce prosodic boundaries after NP1, but rarely after NP2. One possible source of this phrasing in English is the formation of prosodic words that "flout syntactic phrase boundaries, such as *actress who was*" ([3], p. 309). Thus, English has three possible break locations within the complex NP: [the servant of the actress] [who was…]; [the servant] [of the actress who was…]; and [the servant of the] [actress who was…]. There is thus an additional possible pattern for which N2 is grouped with the RC, a prosodic pattern that favors Late Closure. The Late Closure preference in English is then due to the more frequent grouping of NP2 with the RC. In Spanish we would expect more frequent grouping of NP1 and NP2, consistent with Early Closure. Our production study tests these claims.

Productions were annotated using the ToBI (Tones and Break Indices) conventions for English (e.g. [10], cf. [11]) and Spanish (e.g. [12]). The ToBI systems assume two types of tonal targets: localized pitch accents and edge tones, which delimit prosodic phrasal constituents. We focus here on edge tones, which differ for English and Spanish. English ToBI assumes two levels of phrasing, the intermediate phrase (ip), a perceptually smaller break associated with a phrase accent (written as X-) which controls the pitch from the last pitch accent to the end of the phrase; and the intonation phrase (IP), a perceptually bigger break associated with a boundary tone (written as (X%), a local pitch excursion associated with the edge of a phrase. Both phrase accents and boundary tones can

be high or low in pitch. Thus, an ip can end with a L- or H-tone and an IP can end in a L-L%, L-H%, H-L% or H-H% tone combination. Spanish ToBI assumes only one level of phrasing: an intonational phrase. The boundary tones L%, H% and M% are assumed.

2. Production Study

Native English and native Spanish speaking participants read aloud ambiguous sentences of the kind described above and answered comprehension questions gauging their attachment preference for each sentence.

2.1. Participants

Ten native speakers of Midwestern American English and ten native speakers of Mexican Spanish participated in the study. Most English speakers reported knowing one foreign language (either poorly, well, or excellently). Three speakers knew Spanish. Most of the Mexican Spanish speakers report speaking good to excellent English. One speaker grew up bilingually with Spanish and English.

2.2. Method

Participants were recorded reading 34 sentences including 8 ambiguous target sentences containing a complex NP modified by an RC. The Spanish sentences were direct translations of the English sentences. All RCs referred to a person's location (e.g. *who was on the balcony*) to avoid semantic bias towards one interpretation over the other (assuming that it is equally likely for an actress or a servant to be on a balcony). One sentence had to be excluded from the analysis due to a typographical error. Five further tokens were skipped by readers. All filler sentences were structurally ambiguous, but often semantically biased towards one interpretation. All sentences were adapted from [4]. For this study, a total of 67 English and 68 Spanish sentences were analyzed.

The sentences were presented in their entirety on a computer screen. Although not instructed to do so, all participants started reading without scanning or practicing sentences first. After each sentence, participants pushed a button box to see a question that gauged how the sentence had been interpreted. For target sentences, the question was of the form *Who was in/on/at + location?* The response options were NP1 (e.g. *the servant*), NP2 (e.g. *the actress*) or *I don't know*.

All sentences were ToBI annotated by three trained phoneticians. The annotators coded a third of the utterances each and proofed another third of the utterances. Thus, every utterance was looked at by two coders. All utterances that the two coders disagreed on were analyzed by all coders until an agreement was reached. Word durations of V (*shot*), N1 (*servant*) and N2 (*actress*) were measured.

2.3. Attachment Preference

As shown in Figure 1, English readers showed an overall slight low attachment preference (49% vs. 43%; cf. [3]), whereas Spanish speakers showed an overwhelming high attachment preference (78% vs. 19%).
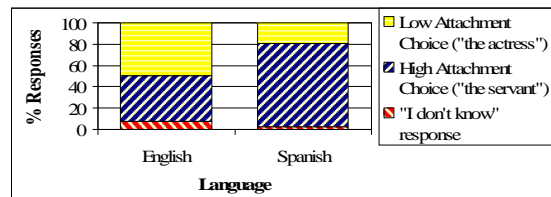

Figure 1: English and Spanish readers' responses to the comprehension questions (e.g. *Who was on the balcony?*).

2.4. Location and Strength of Prosodic Boundaries

If readers' prosody reflects their comprehension, the data in Figure 1 predict prosodic breaks after N2 to be rare in English, but more common in Spanish. Figure 2 shows the prosodic breaks produced for English. The majority of sentences contained no prosodic break after the main verb. At N1, the majority of sentences contained a prosodic break, but only slightly more than a third contained an IP. At N2, on the other hand, IP boundaries were found in the majority of sentences. There is no evidence of prosodic breaks immediately before N2 (not shown). Thus, the reading aloud pattern for these sentences is inconsistent with the (slight) low attachment preference found in sentence comprehension.
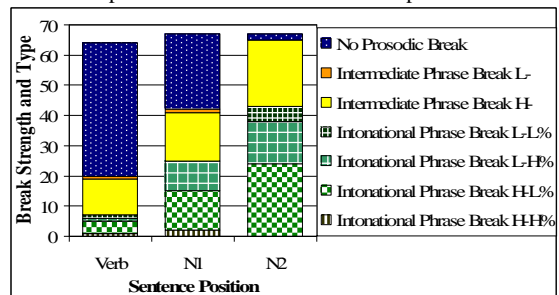

Figure 2: Boundary strength and type at the main verb (V), N1, and N2 of the English utterances.

Figure 3 shows the analysis of prosodic breaks for Spanish. Most boundaries occurred after N2, and boundaries after V and N1 were rare. Compared to English (counting both ips and IPs), there are equally many breaks after N2, but fewer breaks after V and especially after N1.

Thus the findings show no difference in break strength between English and Spanish at N2 and are incompatible with the assumption that the English low attachment preference is due to rare prosodic boundaries after N2, while the Spanish high attachment preference is a result of more frequent breaks at N2.
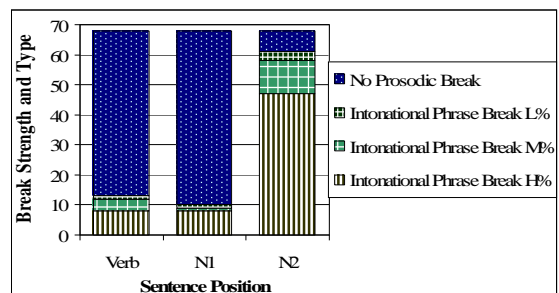

Figure 3: Boundary strength and type at the main verb (V), N1, and N2 of the Spanish utterances.

Previous work on the IPH has also focused on the relative strengths of boundaries following NP1 and NP2. A stronger boundary at N1 is consistent with low attachment, while a stronger boundary at N2 corresponds to high attachment. To investigate whether attachment choices in comprehension matched the prosodies produced when we take into account break strength, we coded all utterances for location of the strongest break. The coding scheme assumed IP were stronger than ip breaks, and lower tones stronger than higher tones. (Low tones tend to indicate greater phrased finality in English (cf. [9]).) Thus the strength codes for English were: No Prosodic Break < H- < L- < L-H% < H-L% < H-H% < L-L%[1]. The codes for Spanish were: No Prosodic Break < H% < M% < L%. Utterances were coded as Strongest Break Follows (SBF) N1 if the break after N1 was stronger than that after N2 and vice versa. If the breaks after N1 and N2 were equally strong, the utterance was coded as Equal Breaks.

Figures 4 (English) and 5 (Spanish) reveal no one-to-one correspondence between produced prosody and sentence interpretation. In fact, all prosodic break patterns show a comparable number of low and high attachment choices in comprehension. The main difference between the English and Spanish data was in the overall number of low and high attachment decisions. This suggests that the prosodic break patterns that speakers produced did not influence their attachment choices.
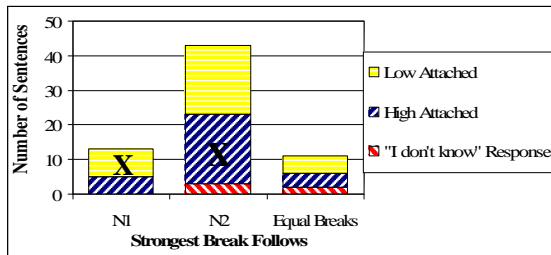

Figure 4: Comparison of sentence comprehension and production for English. The X indicates the productions most consistent with readers' comprehension choices.
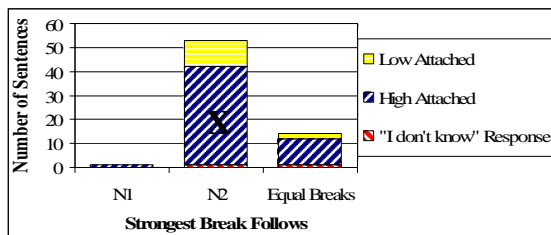

Figure 5: Comparison of sentence comprehension and production for Spanish. The X indicates the productions most consistent with the readers' comprehension.

IPH researchers often assume that readers produce a prosodic break after N1 or after N2 but not after both N1 and N2. To investigate whether the presence or absence of a prosodic boundary at one possible boundary location influenced the likelihood of a boundary at the following possible location, we graphed the break strengths of N1 and N2 as a function of the break strengths of the main verb and N1, respectively. The results are shown in Figures 6 and 7 for English and Figures 8 and 9 for Spanish.

As shown in the graphs, there is no evidence that the presence of a prosodic break at the main verb or N1 led to the absence of a prosodic break at N1 or N2, respectively, or vice versa. Rather, the break strengths at a following location are relatively equally distributed among the break strengths at a preceding location.
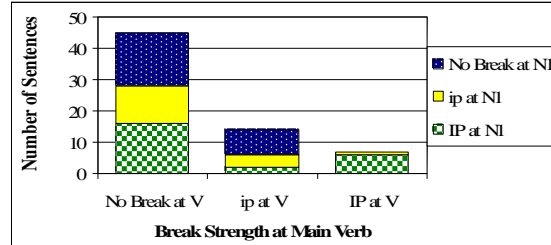

Figure 6: Break strength at N1 as a function of break strength at the main verb for English utterances.
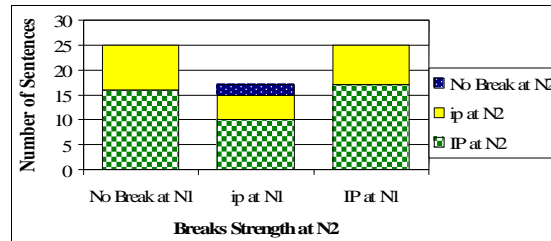

Figure 7: Break strength at N2 as a function of break strength at N1 for English utterances.
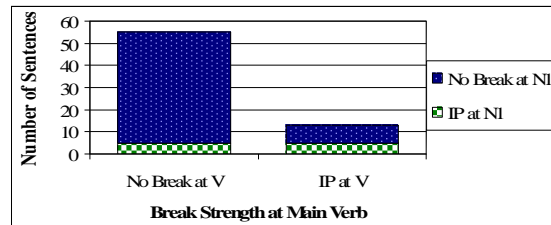

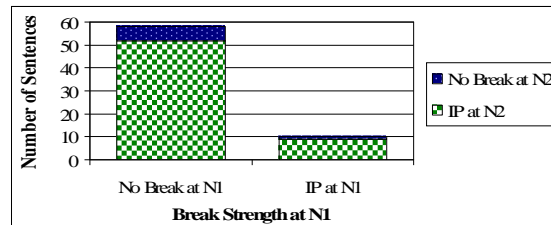Figure 8: Break strength at N1 as a function of break strength at the main verb for Spanish utterances.

Figure 9: Break strength at N2 as a function of break strength at N1 for Spanish utterances.

This result was confirmed by duration measurements of the main verb, N1 and N2, which all show a positive correlation suggesting that as the main verb gets longer, so does N1, and as N1 gets longer so does N2. If the presence of a prosodic boundary at one location were correlated with the absence of a break at a neighboring location, we would have expected the duration of the main verb and N1 as well as N1 and N2 to be inversely related.

---

[1] Since both H-H% and L-L% are utterance-final patterns, they were considered the strongest breaks.

3. Discussion

Our data provides insight into the process of reading texts aloud. For both languages, the likelihood of a prosodic break at a given location depended on both when the last break was uttered and syntactic structure. Overall, readers showed a tendency to prosodically separate the main clause from the relative clause. In addition, some readers produced more frequent prosodic breaks than others, especially at V and N1. Thus, individual speaker differences best account for the variation noted beyond the preference for placing a break before the RC.

Interpretation of the sentence was not correlated with prosodic patterns. This makes sense considering the task: participants had to build an interpretation as they were reading aloud. That is, the message incrementally followed production. It is thus not surprising that readers produced a "default" prosodic pattern, favoring a break before the RC over other locations, rather than basing the prosodic pattern on any interpretation (which had probably not been established yet). Another possibility is that readers could preview the relativiser *who*, which indicates a big upcoming propositional chunk. Producing a prosodic break before the RC may have allowed readers to wrap up preceding information and prepare for the upcoming larger proposition.

Compare this to spontaneous speech, where the message incrementally precedes production. That is, an utterance is produced to convey the message that the speaker intends. With a message already in mind, prosodic patterns are more likely to match the intended message. Furthermore, converting a thought into an utterance is a different task from converting written words into an utterance. Therefore, our data do *not* suggest that prosodic break patterns in natural speech would not correlate with speakers' intended meaning.

It seems that meaning is "larger" than prosody in that many factors contribute to a sentence's meaning, only one of which is prosody. Thus, a given meaning often leads to meaning-consistent prosody, but a given prosody does not necessarily lead to a prosody-consistent meaning since other factors such as word meaning, syntax, and world knowledge further influence how a globally ambiguous sentence is interpreted.

Our data provides evidence against a prosodic account of the differences in overall attachment preference of the NP1 NP2 RC construction across languages. Both low attaching English and high attaching Spanish showed similar production patterns in reading aloud, with most prosodic breaks occurring at N2. Previous production studies in Korean and Japanese ([13], [14]), in which constituent order for the same construction is RC NP2 NP1, report the strongest prosodic boundary after the RC and an overall preference for high attachment choices in both languages. Thus, in languages for which production data similar to ours is available, the preferred strongest break location is between the RC and the complex NP. Thus, the languages analyzed so far prefer a phrasing pattern that is more consistent with a high attachment preference. It seems then that the phrasing pattern across languages is more consistent than the attachment pattern.

4. Conclusions

Our data suggest that attachment preferences in reading the NP1 NP2 RC construction across languages do not differ as a result of the languages' prosodies. We propose that this is the case because message construction incrementally follows production both in silent and out loud reading tasks.

Further studies are needed to investigate which factors influence meaning generation in silent and out loud reading. Finally, more research is needed to better understand the process of reading aloud and how it differs from spontaneous speech and potentially from silent reading.

5. References

[1] Frazier, L., 1978. *On comprehending sentences: Syntactic parsing strategies.* University of Connecticut: Unpublished Doctoral Dissertation.

[2] Fodor, J. D.; Frazier, L., 1980. Is the human sentence processing mechanism an ATN? *Cognition* 8(3), 417-459.

[3] Fodor, J. D., 1998. Learning to Parse? *Journal of Psycholinguistic Research* 27(2), 285-319.

[4] Cuetos, F.; Mitchell, D. C., 1988. Crosslinguistic differences in Parsing: Restrictions on the use of the Late Closure strategy in Spanish. *Cognition* 30(1), 73-105.

[5] Carreiras, M.; Clifton, C, Jr., 1999. Another word on parsing relative clauses: Eye-tracking evidence from Spanish and English. *Memory & Cognition* 27, 826-833.

[6] Gilboy, E.; Sopena, J.-M.; Clifton, C.; Frazier, L., 1995. Argument structure and association preferences in Spanish and English complex NPs. *Cognition* 54(2), 131-167.

[7] Miyamoto, E., 2001. Table listing experimental results of relative clause attachment preferences, available at http://etm4rc.googlepages.com/table.html.

[8] Fodor, J. D., 2002. Prosodic Disambiguation in Silent Reading, *Proceedings of the North East Linguistic Society,* 32. Amherst: University of Massachusetts, 113-132.

[9] Pierrehumbert, J.; Hirschberg, J., 1990. The meaning of intonational contours in the interpretation of discourse. In *Intensions in Communication,* P. R. Cohen; J. Morgan; M. E. Pollack (eds.). Cambridge, MA: MIT Press, 271-311.

[10] MIT OpenCourseWare, 2006. *Transcribing Prosodic Structure of Spoken Utterances with ToBI*, available at http://ocw.mit.edu/OcwWeb/Electrical-Engineering-and-Computer-Science/6-911January--IAP--2006/CourseHome/index.htm.

[11] Beckman, M.; Pierrehumbert, J., 1986. Intonational Structure in Japanese and English. *Phonology Yearbook* 3, 255-309.

[12] Beckman, M.; Diaz-Campos, M.; McGory, J.; Morgan, T., 2002. Intonation across Spanish, in the Tones and Break Indices Framework. *Probus* 14, 9-36.

[13] Jun, S.-A.; Koike, C., 2003. Default Prosody and RC Attachment in Japanese. *Talk given at the 13th meeting of Japanese-Korean Linguistics conference*, Tucson, Arizona. [paper to appear in Japanese-Korean Linguistics 13, Stanford: CSLI.]

[14] Jun, S.-A., 2004. The Intermediate Phrase in Korean Intonation: Evidence from Sentence Processing. *Talk presented at the International Conference on Tone and Intonation in Europe*, Santorini, Greece. [paper to appear as a chapter in Tones and Tunes: Studies in Word and Sentence Prosody, eds. by C. Gussenhoven and T. Riad. Berlin: Mouton de Gruyter.]