PROCESSAMENTO DE RELATIVAS EM PORTUGUÊS BRASILEIRO

Edson T. Miyamoto1 & Ingrid Finger2

RESUMO: Como as pessoas lêem palavras uma a uma e extraem o significado de uma sentença? Essa é a questão básica que se tenta responder quando se pesquisa o processamento de sentenças. Um dos pontos de interesse nessa área é o processamento de orações relativas. O presente artigo reporta um estudo de questionário conduzido em Porto Alegre e São Paulo com itens traduzidos do inglês (Carreiras & Clifton, 1993; Cuetos & Mitchell, 1988) que visa investigar em que medida as variações regionais do português brasileiro contribuem para a solução de algumas questões teóricas pendentes.

PALAVRAS-CHAVE: processamento de sentenças, orações relativas, variação dialetal

INTRODUCTION

In the field of sentence processing, the primary goal is to investigate how people read (or hear) sentences word by word and understand their meaning.

Assume that sentence processing can be characterized by the following two components: a database that contains knowledge about a given language (e.g., its grammar, social conventions, etc), and an algorithm that uses the database in order to process sentences.

Clearly, the database is different for different languages given that grammars and social conventions vary. But, is it possible that a single algorithm is used to process all languages in the world? In other words, is it possible that all parametric variations are restricted to the database as implied by the Universal Processor Hypothesis?

(1) UPH (Universal Processor Hypothesis): All human languages are processed by the same algorithm.

The UPH is often assumed these days either tacitly or explicitly. Therefore, any apparent differences in the processing of two languages have to be explained in terms of interactions between the distinct databases of those languages and the single parsing algorithm. Such interactions have been extensively investigated in order to explain why locality preferences are violated in the processing of relative clauses in some languages (Cuetos & Mitchell, 1988).

1.1 LOCALITY PREFERENCES IN Parsing

At least since the 1970’s it has been known that readers preferentially attach ambiguous phrases to the closest head available (see Frazier, 1987, for a summary on late closure and its precursors). More recently, it has been suggested that such locality preferences are caused by working memory constraints (see Gibson, 1998, and references therein).

1 University of Tsukuba / Nara Institute of Science and Technology, Japan.
2 Programa de Pós Graduação em Letras, Universidade Católica de Pelotas.
The preference for the local head (or *locality* for short) is observed in various types of constructions across different languages as in the examples below (from Miyamoto, 1999), where the modifier in italics (e.g., *with binoculars* in (2)) is preferentially attached to the closest of the two underlined heads (*looking* rather than *saw*).

(2) a. The policeman saw the man who was looking at the woman *with binoculars*.  
   b. Policajt viděl muže který koukal na ženu *s dalekohledem*. (Czech)

(3) a. The daughter of the professor in Sudan likes apples.  
   b. Die Tochter des Professors *in Sudan* mag Äpfel. (German)

(4) a. I ate the ice-cream that I bought yesterday.  
   b. *Kinou katta aisu-kurimu-o tabeta*. (Japanese)

(5) a. John bought a laptop computer with a TFT screen and a fax machine.  
   b. Jo *a comprou um laptop com um monitor TFT e um fax*.

1. RELATIVE CLAUSE ATTACHMENT

Although locality is believed to be a universal characteristic of the human processing algorithm, it is not obeyed in the attachment of ambiguous relative clauses (RCs) as in (6) (from Cuetos & Mitchell, 1988), therefore challenging the validity of the UPH.

(6) a. Someone shot the servant of the actress [RC *who was on the balcony*].  
   b. Alguien disparó contra el criado de la *actriz* [RC *que estaba en el balcón*]. (Spanish)

Despite the fact that a preference for associating the RC to *actress* in (6a) (i.e., *the actress was on the balcony*) has been found in English, a non-local preference (the attachment of the RC to *criado* in (6b)) was found in Spanish (Cuetos & Mitchell, 1988). Non-local attachment preferences have also been observed in several other languages tested thus far (e.g., Dutch, French, German, Japanese; for a more extensive list of languages tested, see http://www.lingua.tsukuba.ac.jp/etm/rc).

2. BRAZILIAN PORTUGUESE

Because the results regarding RC interpretation preferences in Brazilian Portuguese (BP) have been contradictory, we investigated the extent to which regional variations in BP present differences in RC processing and thus potentially contribute to the solution of theoretical questions in the area.

2.1. WORD ORDER

Various grammatical parameterizations have been singled out in order to explain the differences in RC attachment preferences. Among those, rigid word order has been
claimed to be responsible for the low attachment preference in English (Gibson et al., 1996). However, because word order in French is relatively rigid and RC attachment preference is nevertheless non-local in this language (Zagar, Pynte & Rativeau, 1997), a more fine-grained definition of rigid word order is necessary.

For present purposes, we define a language as having rigid word order if it does not allow an adverb to intervene between a transitive verb and a direct object (following Miyamoto, 1999; see Pollock, 1989, for the syntactic consequences of such differences between English and French). Word order flexibility as attested by adverb intervention may correlate with non-local RC attachment preferences because locality is more easily violated in languages with flexible word order.

(7) **Hypothesis**: in languages with rigid word order (i.e., languages that do not allow an adverb to intervene between a transitive verb and its direct object), readers prefer to attach ambiguous RCs to the most local noun in constructions such as (6).

The hypothesis above correctly predicts the local attachment preference in English as opposed to the non-local preference in Dutch, French, German and Spanish. BP is an ideal language to further test the hypothesis.

2.2. WORD ORDER IN BRAZILIAN PORTUGUESE

Linear word order in BP with respect to adverb placement is a controversial issue. There is no theoretical consensus as to whether BP can be considered a verb-raising language, and to what extent adverb placement is a reliable test for verb-movement. We will assume in the following that BP – at least, in some of its dialects – is similar to English in that grammaticality is degraded when an adverb intervenes between a verb and its object (Rohrbacher, 1994, 1995, and references therein). We support this observation with the following judgements, which seem to hold in the BP spoken in São Paulo.

(8) a. *? João abraça frequentemente a Maria.
   b. João abraça a Maria frequentemente.
   c. João frequentemente abraça a Maria.
   “João frequently hugs Mary.”

(9) a. *? João perde sempre a carteira.
   b. João sempre perde a carteira.
   “João always misplaces his wallet.”

Both (8a) and (9a) need emphasis on a Maria and sempre respectively in order to sound natural in BP. Despite the possible regional differences regarding grammaticality judgements, the ungrammaticality of sentences such as (10) below, in which the article a (“the”) has been removed and emphasis on the object NP is more difficult, seems to support the claim that adverb placement between verb and direct object degrades grammaticality.

(10) * João abraça frequentemente Maria.
The markedness of adverbs between verb and object in BP has been related to the loss of second person agreement (Rohrbacher, 1994, 1995). Consequently, regions in Brazil where the second person pronoun inflection is still in use should present grammars with more flexible word order and, hence a stronger preference to attach the RC to the non-local noun. We can thus build a table correlating the three phenomena as follows, where the last column are the predictions made by hypothesis (7).

Table 1 – Predictions of RC attachment on the basis of person agreement

<table>
<thead>
<tr>
<th></th>
<th>second person agreement</th>
<th>intervening adverb</th>
<th>RC attachment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Porto Alegre</td>
<td>Yes</td>
<td>Yes?</td>
<td>Non-local</td>
</tr>
<tr>
<td>São Paulo</td>
<td>No</td>
<td>No</td>
<td>Local</td>
</tr>
</tbody>
</table>

3. PREVIOUS RESULTS IN BRAZILIAN PORTUGUESE

As summarized below, previous results in RC attachment in BP have been contradictory.

3.1. A LOCAL ATTACHMENT RESULT

Miyamoto (1999) reported a local attachment preference for BP using a self-paced reading task with sentences disambiguated by number agreement like the following.

(11) a. Non-local attachment
Uma aluna criticou os professores do curso [RC que foram escolhidos pelos estudantes].
“A student criticized the teachers of the course that were chosen by the students.”

b. Local attachment
Uma aluna criticou o professor dos cursos [RC que foram escolhidos pelos estudantes].
“A student criticized the teacher of the courses that were chosen by the students.”

Reading times were slower at the RC if it had to attach to the non-local noun as in (11a). However, the study had a number of problems. First, the sentences were considerably different from the original sentences used in the English/Spanish study (Cuetos & Mitchell, 1988). Second, the BP speakers were residents in the United States, therefore it is conceivable that knowledge of English interfered with attachment preferences (but see Gibson, Pearlmutter & Torrens, 1999, for a non-local attachment result with Spanish speakers living in the United States). Third, most problematic is that recent studies suggest that number agreement can affect the reading times of RC attachment experiments (Carreiras, Betancort & Meseguer, 2001). Moreover, an off-line questionnaire using ambiguous items like the following did not reveal any statistically reliable preference.
Ambiguous attachment
Uma aluna criticou o professor do curso que foi escolhido pelos estudantes.
“A student criticized the teacher of the course that was chosen by the students.”

3.2. NON-LOCAL ATTACHMENT RESULTS

Off-line experiments conducted in Rio de Janeiro (Maia & Maia, 2001) using a subset of the items from Cuetos & Mitchell (1988) translated into BP, and in Porto Alegre with a new set of items (Finger & Zimmer, 2002) suggest that RC attachment preference is in fact non-local in BP. The present study attempts to address this controversy by conducting an experiment in two different regions of the country, namely São Paulo (SP for short) and Porto Alegre (POA).

4. EXPERIMENT

4.1. PARTICIPANTS

A total of 61 undergraduate students in engineering or computer science based in POA volunteered to participate in the experiment. In SP, 14 undergraduates in computer science have been tested so far. At the time of data collection, all participants were over 18 years old and reported to be monolingual native speakers of BP.

4.2. MATERIALS

4.2.1. QUESTIONNAIRE 1

Questionnaire 1 contained 24 items based on materials from Cuetos & Mitchell (1988). A total of 30 participants were tested in PA and 14 were tested in SP. The 24 test items were presented interspersed with 56 unrelated filler items in pseudo-random order so that at least one filler intervened between two consecutive test items. Sentence (13) below is an example sentence followed by the forced choice task.

(13) Alguém atirou no empregado da atriz que estava na sacada.
Quem estava na sacada? [   ] o empregado [   ] a atriz

4.2.2. QUESTIONNAIRE 2

Questionnaire 2 was composed of 16 items based on materials from Carreiras & Clifton (1993) and 6 items from Cuetos & Mitchell (1988), totaling 22 target sentences. Here again the test items and the 62 filler items were presented in pseudo-random order. Sentence (14) is an example of a test item.

(14) O policial prendeu a irmã da criada que está grávida.
Quem está grávida? [   ] a irmã [   ] a criada
In both questionnaires the order of the alternatives was counterbalanced creating two versions of the task for each test item. For (14), the second version of the task would be as follows.

(15) Quem está grávida? [ ] a criada [ ] a irmã

For each questionnaire, the two versions of each item were distributed into two lists following a Latin Square design. Each participant saw only one list interspersed with filler items.

4.3. RESULTS

Mean and standard error for the non-local attachment preference for each city (POA or SP) and each questionnaire (1 or 2) are reported in Table 2.

<table>
<thead>
<tr>
<th>City</th>
<th>Quest.</th>
<th>n</th>
<th>Mean % (Standard error)</th>
<th>Breakdown by Participants (non-local/local)</th>
<th>Breakdown by Items (non-local/local)</th>
</tr>
</thead>
<tbody>
<tr>
<td>POA</td>
<td>1</td>
<td>30</td>
<td>71 (2.7)</td>
<td>25/3</td>
<td>22/2</td>
</tr>
<tr>
<td>POA</td>
<td>2</td>
<td>31</td>
<td>74 (2.6)</td>
<td>29/1</td>
<td>19/3</td>
</tr>
<tr>
<td>SP</td>
<td>1</td>
<td>14</td>
<td>62 (5.0)</td>
<td>10/3</td>
<td>22/2</td>
</tr>
</tbody>
</table>

In questionnaire 1, the POA participants showed a non-local attachment preference overall (71%). Out of the 30 participants, 25 chose non-local attachment more often than local attachment, three chose local attachment more often than non-local attachment, and the remaining two participants had equal numbers of local and non-local choices. Among the 24 items, 22 were more often interpreted with RCs attached non-locally in contrast to two items that were interpreted locally more often.

The SP participants also preferred the non-local attachment interpretation overall (62%). Out of the 14 participants, ten chose non-local attachment more often than local attachment, three chose local attachment more often than non-local attachment, and for one participant the numbers of local and non-local choices were the same. In the breakdown by items, the results were similar to what was reported above for the POA participants.

The difference between the attachment preferences of the participants from POA and from SP (lines 1 and 3 of Table 2) was reliable in the analysis by items (F2(1,23)=8.79, p<0.01) but not in the analysis by participants (F1(1,42)=2.14, p=0.15). The lack of significance in the participant analysis may be due to the small number of participants from SP, we are in the process of collecting more data in order to further investigate potential regional differences.

For questionnaire 2, non-local attachment was preferred 74% of the time. The breakdown by participants was as follows: 29 participants chose non-local attachment more often than local attachment, one participant preferred local attachment overall, and one participant’s choices were split evenly. Among the 22 target items tested, 19 had their RCs interpreted non-locally more often than locally, in contrast to three items that were interpreted with RCs attached locally more often.
CONCLUSION

Similar populations in two cities judged the non-local attachment interpretation more natural. Thus, word order flexibility does not seem to be a good predictor of RC attachment. However, there is a tendency for the population in SP to have a weaker preference for non-local attachment, therefore a weaker version of (7) may still hold. More data have to be collected in SP and a more detailed comparison should be conducted. Further results are necessary in order to explore other regional differences and alternative explanations.

REFERENCES


ABSTRACT: A basic question investigated in the field of sentence processing is how people read sentences word by word and understand their meaning. The present article reports the results of a questionnaire study that was conducted in Porto Alegre and São Paulo with items translated from English (Carreiras & Clifton, 1993; Cuetos & Mitchell, 1988) in order to investigate the extent to which regional variations in Brazilian Portuguese can contribute to the solution of outstanding questions in the processing of ambiguous relative clauses.

PALAVRAS-CHAVE: sentence processing, relative clauses, regional variation