

## THE RESOLUTION OF AMBIGUOUS ANAPHORA IN PORTUGUESE, ENGLISH AND SPANISH (A PILOT STUDY)

Amanda Maraschin BRUSCATO\*

Jorge BAPTISTA \*\*

- **ABSTRACT:** This paper presents the results of a pilot study that aims to analyse the differences in the resolution of ambiguous anaphora in Portuguese, English, and Spanish as first languages (L1) and as second languages (L2). To collect the data, a questionnaire was developed and applied to national and foreign students at the University of Algarve, in Faro, Portugal. The conclusions are that native English speakers usually interpret the anaphoric pronoun in the subject position of the subordinate clause as corresponding to the subject of the main clause, regardless of whether the pronoun receives emphasis or not. Portuguese and Spanish native speakers, on the other hand, seem to use the Position of Antecedent Strategy (CARMINATI, 2002) in cases of anaphora, but the preference becomes less evident when reversing the order of sentences (cases of cataphora). Besides the syntactic variables, the phrases were designed to investigate the influence of the semantic relation between the verbs to the resolution of ambiguous anaphora.
- **KEYWORDS:** Ambiguity resolution; anaphora; English; Portuguese; Spanish.

### Introduction

This paper aims to analyse the differences in the resolution of ambiguous anaphora in Portuguese, English, and Spanish as first language (L1) and as second language (L2). To obtain the data, a questionnaire was elaborated and applied to national and foreign students at the University of Algarve, in Faro, Portugal.

The term *anaphora* has its origin in the Greek word *anaphorá*, which means “repetition” (NASCENTES, 1955, p.28). Anaphora is a textual reference device in which an anaphor takes up the mention of a given entity/object (the antecedent) that appeared earlier in the discourse. According to Halliday and Hasan (1976), anaphora is an important mechanism of textual cohesion that, by taking up elements in the

---

\* Universidade do Algarve (UALG). Faro - Algarve, Portugal. ORCID: 0000-0003-0660-9098. amandabruscato@gmail.com.

\*\* Universidade do Algarve (UALG). Faro - Algarve, Portugal. ORCID: 0000-0002-3259-5550. jrbaptis@ualg.pt.

text, allows the discourse to progress. This mechanism can be illustrated as in the example (1)<sup>1</sup>:

(1) *Annai hugged and Øi kissed Paulj when hej asked to marry heri.*

In (1), two types of anaphora are observed: *pronominal anaphora* and *zero anaphora*. Pronominal anaphora occurs when a pronoun is used to retrieve an antecedent, such as the use of the pronoun *he* to retrieve the name *Paul* and the pronoun *her* to retrieve the name *Anna*. Zero anaphora, on the other hand, occurs when the ellipsis (or *reduction to zero*) is used instead of repeating the antecedent to retrieve it; this is marked in the example with the symbol *Ø*.

Anaphora can also be classified as *coreferential*, when it takes up the mention of the same entity in the text, as in (1); or *non-coreferential*, when it relies on the interpretation of previous terms to introduce the mention of a new entity, as seen in (2).

(2) *They say that having a peti is good for your health. When I adopted my catj, I was very happy.*

Finally, anaphora, understood as the reference relationship established between the elements of the text, can either take up previously introduced antecedents or anticipate them, as seen respectively in (3) and (4):

(3) *Anna kissed Pauli when hei woke up.*

(4) *When hei woke up, Anna kissed Pauli.*

If the antecedent appears in the text before the anaphor, as in (3), we identify a case of *forward anaphora* or *anaphora* properly speaking. If, in the order in which the words are presented in the text, the anaphor appears before its antecedent, as in (4), we identify a case of *backwards anaphora* or just *cataphora*.

In this article, we use the word *anaphora* to refer to both anaphora and cataphora, considering that it is the same mechanism, regardless of the relative position of the anaphor and its “antecedent”. When it is necessary to distinguish anaphora and cataphora, we will use both terms.

Two types of anaphora are considered in this research: *pronominal anaphora* and *zero anaphora*. The aim is to investigate how native speakers and learners of Portuguese, English, and Spanish resolve ambiguous anaphora in these languages.

Initially, a brief review of the literature on the subject will be carried out, to better frame this study in the state of the art. Then, the methods adopted in the research will

---

<sup>1</sup> In this work, the indices (*i, j, k*) will be used to indicate the coreference between the anaphors and their antecedents. By convention, the indices *i* and *j* will be used for the entities mentioned in the sentence and the index *k* will be used when the anaphor refers to an entity other than those mentioned in the sentence.

be described. Finally, the results obtained will be presented and discussed, pointing out clues for future work.

## Literature review

The problem of anaphoric ambiguity resolution, i.e. determining the antecedent of an anaphor when it is ambiguous, is not a new problem in Linguistics, and there is extensive literature on the subject. In this review, some of the studies relevant to the investigation will be presented.

In their research on pronominal anaphora resolution, Crawley, Stevenson, and Kleinman (1990) investigated which strategy English speakers would use, considering two alternative strategies: (i) either the pronoun would refer to the term occupying the subject position of the previous clause, or (ii) it would refer to the term occupying a syntactic function parallel theirs in the previous clause. Consider the following examples:

- (5) *Johni hit Billj and hei,j ran away.*
- (6) *Johni hit Billj and Mary kicked himi,j.*

Using both strategies, the pronoun *he* in example (5) would be an anaphor of the noun *John*. In example (6), however, according to the first strategy, the pronoun *him* would be an anaphor of the noun *John*, while, according to the second, it would refer to *Bill*.

To answer which anaphora resolution strategy would be used by speakers in cases such as example (6), the authors asked 132 university students to read 40 texts and answer which of the antecedents the pronoun referred to. The texts were composed of three sentences, as shown in (7):

- (7) *Brenda and Harriet were starring in the local musical. Bill was in it too and none of them were very sure of their lines or the dance steps. Brenda*i* copied Harriet*j* and Bill watched her*i,j*.*

In each of the 40 texts, the first sentence introduced two referents; the second introduced a third; and the third took them up through three names and presented a pronoun in the complement position of the last clause. Two versions were created for each text: an ambiguous version, in which the possible antecedents of the anaphor had the same gender; and an unambiguous version, in which, due to gender restriction, there was only one possible antecedent for each anaphor.

The sample was divided into two groups of 48 and one of 36 participants. The group of 36 individuals read only the ambiguous versions of the texts and selected

the antecedent of the pronoun. The other two groups of 48 individuals read both the ambiguous and the unambiguous versions. These two groups were necessary so that no participant could read both versions of the same text.

According to the authors, three evaluators analysed the 40 texts so that there was no influence of general knowledge on pronominal anaphora resolution. The only explicit evaluation criterion, however, was the plausibility of the pronoun referring to both the subject and complement of the previous clause.

In unambiguous cases, due to gender restriction, participants did not use either strategy on anaphora resolution. However, in the ambiguous cases, they showed a preference for the first strategy, i.e., the interpretation of the pronoun as coreferential to the subject of the previous clause.

Anaphora resolution strategies can vary depending on the language. According to the Null Subject Parameter (CHOMSKY, 1981; RIZZI, 1982), English is considered a mandatory subject language, unlike Spanish and Portuguese, which allow the ellipsis of the subject. Anaphoric subjects in these Romance languages, therefore, can be expressed either by pronominal anaphora or zero anaphora. Let us compare the following sentences:

(8a) *Mariai hace lo que ellak quiere.*

(8b) *Mariai hace lo que Øi quiere.*

(9a) *Maryi does what shei,k wants.*

(9b) *\*Mary does what Ø wants.*

While in Spanish, sentences as (8a) and (8b) are considered grammatical, in English (9a) is grammatical, but not (9b). Now let us go back to example (5):

(5) *Johni hit Billj and hei,j ran away.*

Although English speakers prefer to solve pronominal anaphora in sentences as (5) by interpreting the antecedent of the pronoun as the subject of the first clause, the fact is that there is ambiguity, since the sentence would be the same for the interpretation of the antecedent as the complement of the previous clause. In null subject languages, however, research has shown that speakers tend to use and interpret zero anaphora in subject position as coreferential to an antecedent also in the subject position, case of (8b), and the pronoun in the subject position as coreferential to another entity, case of (8a).

This preference, named *Position of Antecedent Strategy*, was identified by Carminati (2002) in relation to Italian intra-sentential anaphora. Her experiment was replicated in Spanish (ALONSO-OVALLE *et al.*, 2002), proving the speakers' preference in using and interpreting zero anaphora in the subject position (instead of pronominal anaphora) as corresponding to the antecedent in subject position in either intra-sentential or inter-sentential, ambiguous or unambiguous contexts.

In European Portuguese, Lobo and Silva (2016) investigated the anaphoric intra-sentential ambiguity in adverbial subordinate clauses with the conjunction *when*. The data obtained from adult participants demonstrated a preference for interpreting the elided subject in the subordinate clauses as coreferential to the subject of the main clauses, and pronominal subject in the subordinate clauses as coreferential to the complement of the main clauses. However, the lexical knowledge of verbs was not considered as a possible influence. According to the authors, the anaphoric subject's interpretation would not be influenced by the lexical knowledge of verbs in adverbial clauses.

On the other hand, Morgado (2013) shows that semantic information is relevant for the resolution of ambiguity. The author used for her study verbs that required human subject and complement and analysed how the thematic role attributed to the subject and complement influences anaphora resolution. Her intra-sentential experiment was built using concessive clauses. The author also concluded that there is a preference for the resolution of zero anaphora as coreferential to the subject of the main clause and pronominal anaphora as coreferential to the complement of the main clause.

The research mentioned so far has investigated the preference of native speakers for the resolution of ambiguous anaphora. Below are some studies comparing the preference of native and non-native speakers for the resolution of ambiguity.

Comparing two experimental settings where Italian is either L1 or L2, Sorace and Filiaci (2006) investigated pronominal and zero anaphora and cataphora in intra-sentential contexts. The group that spoke Italian as L1 was formed by 20 adults, while the group that spoke Italian as L2 was formed by 14 English native adults. This group started studying Italian after puberty, had a high level of proficiency in the language and had been living in Italy for at least 18 months.

For the research, the authors adapted 20 sentences that they had created in a previous study with Tsimpli and Heycock (TSIMPLI *et al.*, 2004). However, neither article explains the criteria for the choice of verbs and conjunctions, they only explicitly state that all the sentences were formed by a main clause that had an animated subject, a transitive verb, and an animated complement; and by a subordinate clause that had a pronoun or ellipsis, a verb and a complement. In the example they presented, the adverbial subordinate clause precedes the main clause and uses the temporal conjunction *while* is used, as seen in (10):

(10) *While she<sub>i,j,k</sub> wears the coat, the mother<sub>i</sub> gives a kiss to the daughter<sub>j</sub>.*

Therefore, 5 sentences with pronominal anaphora, 5 with pronominal cataphora, 5 with zero anaphora and 5 with zero cataphora were used. For each sentence, there were three images that illustrated the different interpretations. After reading each sentence, the participants selected the image that represented their interpretation. The authors found that both native and non-native Italian speakers presented very similar results in

zero anaphora and zero cataphora resolution, but results were different in pronominal anaphora and pronominal cataphora resolution, as shown in the following table:

**Table 1** – Results of Sorace and Filiaci (%)

|           |            | Natives |            |       | Non-natives |            |       |
|-----------|------------|---------|------------|-------|-------------|------------|-------|
|           |            | Subject | Complement | Other | Subject     | Complement | Other |
| Anaphora  | Pronominal | 8       | 82         | 11    | 27          | 60         | 13    |
|           | Zero       | 51      | 44         | 5     | 46          | 43         | 11    |
| Cataphora | Pronominal | 12      | 24         | 64    | 47          | 25         | 28    |
|           | Zero       | 85      | 11         | 4     | 85          | 9          | 6     |

**Source:** Authors' elaboration based on Sorace and Filiaci (2006).

It is evident that non-native speakers have related the pronoun to the subject more often than native speakers, possibly due to the preference in English for interpreting the pronoun in the subject position of the subordinate clause as coreferential to the subject of the main clause. The main difference, however, was the preference of Italian native speakers for solving pronominal cataphora by relating the pronoun to another referent, i.e. neither to the subject nor to the complement of the main clause.

According to the authors, the violation of the Position of Antecedent Strategy, unlike in cases with zero anaphora, is acceptable for the resolution of pronominal anaphora in unambiguous sentences like (11):

(11) *Johni hugged Maryj when hei saw herj.*

Jegerski, VanPatten, and Keating (2011) suggested that the Position of Antecedent Strategy is not applicable to Spanish pronominal anaphora. They investigated the resolution of ambiguous intra-sentential anaphora in both English and Spanish as a first language (L1) and in Spanish as a foreign language (L2). Twenty sentences with anaphoric ambiguity were used in English and twice as many in Spanish, since in addition to pronominal anaphora, Spanish also allows zero anaphora. All the examples presented two sentences connected by some temporal conjunction: *while*, *when*, *before* or *after*.

The authors confirmed that most Spanish native speakers interpret the ellipsis of the subject as coreferential to the previous subject (72.83%). The interpretation of pronominal anaphora in Spanish as L1, however, did not agree with the Position of Antecedent Strategy. According to the results obtained, about half of the pronouns (52.39%) were related to the antecedent in subject position of the main clauses, while the other half (47.61%) were related to the antecedent in complement position.

Regarding speakers of Spanish as a foreign language, participants were divided according to intermediate and advanced proficiency level. The two groups presented

similar results to each other and slightly different from those obtained with native speakers. Probably due to the influence of their first language (English), participants interpreted zero anaphora as coreferential to the subject in 60.68% of the cases, slightly less than native speakers, and they interpreted the pronominal anaphora as coreferential to the subject in 55.23% of the cases, slightly more than native speakers.

Also aiming to compare the resolution of ambiguous anaphora in English and Spanish as L1 and L2, Valenzuela, Licerias and Morelos (2011) carried out two studies in which they used 24 sentences with subordinate clauses, half of which half featured cataphora, while the other half presented anaphora. In Spanish, half of the cases of anaphoras/cataphoras were elliptical and the other half were pronominal. In English, half of the sentences had an unstressed pronoun, and the other half showed a stressed pronoun (which was represented by capital letters), as shown in (12a) and (12b), respectively:

(12a) *Alexi bumped into Johnj while hei,j was riding hisi,j bike.*

(12b) *Soniai saw Marisaj while SHEi,j was brushing heri,j teeth.*

For both languages, native speakers and learners were asked to resolve cases of ambiguity regarding the subject of the subordinate clauses. The data regarding the answers about who was the subject of the clauses in English are organised in Table 2 below:

**Table 2** – Results of Valenzuela, Licerias, and Morelos (%)

|                           | Natives |            |      | Non-natives |            |      |
|---------------------------|---------|------------|------|-------------|------------|------|
|                           | Subject | Complement | None | Subject     | Complement | None |
| Anaphora without emphasis | 49      | 49         | 2    | 64          | 32         | 4    |
| Anaphora with emphasis    | 41      | 59         | 0    | 40          | 50         | 10   |

**Source:** Authors' elaboration based on Valenzuela, Licerias, and Morelos (2011).

It was noticed that native English speakers, when answering who the subject of the subordinate clauses was, were not influenced by the emphasis on pronouns. On the other hand, English learners, who were Spanish native speakers, interpreted non-emphatic anaphora in English as they interpret zero anaphora in Spanish, relating it to the subject of the main clause; they also interpret emphatic anaphora in English as they would with pronominal anaphora in Spanish, relating it to the complement.

Except for Morgado (2013), the studies discussed in this section did not indicate any semantic criteria for the choice of verbs used in the surveys. Moreover, they presented different results regarding the resolution of ambiguous anaphora. To investigate ambiguity resolution in Portuguese, English, and Spanish, it was decided to better

control in this research the criteria for sentence elaboration and to compare data obtained from both native speakers and learners of these three languages.

## Methods

Inspired by the study of Valenzuela, Licerias, and Morelos (2011), we prepared a questionnaire with ambiguous sentences in which, for each sentence, the students were asked to indicate the antecedent of the subject of the subordinate clause. Three alternatives were always presented:

- (i) the subject of the main clause;
- (ii) the complement of the main clause;
- (iii) another entity, distinct both from the subject and from the complement of the main clause.

The aim of the research was to determine whether there would be differences in strategies for resolving intra-phrasic ambiguity depending on the following variables:

- (i) Portuguese, English or Spanish as L1;
- (ii) Portuguese, English or Spanish as L2;
- (iii) anaphora or cataphora;
- (iv) pronoun or ellipsis (in English, pronoun with or without emphasis);
- (v) influence or not of the relation between the verbs.

The distinction between the complement use or not of emphasis in English pronouns, previously used by Valenzuela, Licerias, and Morelos (2011), was chosen to identify whether English learners interpret the difference between ellipsis and unstressed pronoun in Portuguese as equivalent to the difference between stressed/unstressed pronouns in English. According to Akmajian and Jackendoff (1970), Szwedek (1980), and Luján (1986), the unstressed pronoun in the subject position of the subordinate clauses would be coreferent to the subject of the main clauses, while the use of emphasis on the pronoun would prevent such interpretation. However, the authors made these analysis proposals without having collected data from native speakers. In contrast, Valenzuela, Licerias, and Morelos (2011) concluded with their study that, for native English speakers, stressed pronouns did not entail any significant difference in resolving ambiguity. We then decided to investigate with native English speakers whether the emphatic use of the pronoun would interfere with the identification of its antecedent.

Twenty-three students from the University of Algarve, aged between 20 and 30, participated in this research, most of them female (4/5). The average age was approximately 23 years, with a standard deviation of 3.3. Of the 23 students, 13 were Portuguese speakers attending the fifth semester of the degrees in *Languages*,



*Literatures, and Cultures or Languages and Communication*, being 7 students of Spanish, 10 of English (6 participants studied both English and Spanish) and 2 of French.

Among the participants, five were Spanish native speakers, learners of Portuguese as a foreign language, and five were English native speakers, all students at the Portuguese university and residents in the country for less than a year. The English-speaking group consisted of four women and one man, four from North America and one from New Zealand; and the Spanish-speaking group also consisted of four women and one man, four from Latin America and one from Spain.

Concerning Latin American students, language variety might have constituted an issue for the experimental setting. According to Toribio (2000), the Caribbean variety of Spanish differs from the other varieties in relation to the Null Subject Parameter. In our research, however, there were no speakers of this variety. Despite our initial interest in comparing how the participants would solve ambiguous anaphora in L1 and L2, there were not enough native English speakers studying Portuguese as a foreign language. For this reason, it was not possible to investigate how the ambiguity in Portuguese would be solved by this type of informant.

The applied questionnaire consisted of 16 ambiguous sentences, which had two possible antecedents of the same gender (e.g., *Mateus harassed George when he interviewed him*); and 4 unambiguous sentences, which, due to gender restriction, had only one possible antecedent (e.g., *Diana kissed Daniel when she visited him*). The unambiguous sentences were used as control sentences to see if participants were effectively aware of the questionnaire and committed to its resolution.

The sentences were presented in a random order, so that there was no interference from the researchers in this choice, but it was always the same for all participants, so that the order of presentation did not interfere in the elicitation of responses. Of the 20 sentences prepared for each language, 10 were cases of anaphora and 10 were cases of cataphora. While 5 anaphoras and 5 cataphoras had ellipsis in Portuguese and Spanish and unstressed personal pronoun in English, the other 5 anaphoras and 5 cataphoras had unstressed personal pronoun in Portuguese and Spanish and stressed personal pronoun in English, as shown in the Appendix of this paper.

The verbs used were chosen from a database of lexical-syntactic constructions of European Portuguese verbs (BAPTISTA, 2013). Sixteen unambiguous verbs from this class were selected for the elaboration of the ambiguous sentences. These verbs select a human subject and a human complement, and no further complements. Semantically, these verbs featured a generic agent subject and a direct patient object. Since many of these verbs denote violent acts, 4 verbs from other categories were used for the control sentences to “smooth” the reading of the questionnaire.

All the main sentences were constructed with a person’s name in the subject position, a verb in the simple past and another person’s name in the complement position; in the subordinate clauses, the temporal conjunction *quando/cuando/when* was used along with a 3<sup>rd</sup> person singular subject personal pronoun or its ellipsis, another verb in the simple past and a 3<sup>rd</sup> person singular complement personal pronoun. The sentences

were initially written in Portuguese and then translated into Spanish and English. As stated at the beginning of this section, for each sentence, students were asked who the subject of the subordinate clause referred to, and three alternatives were offered: the name in the subject position of the main clause, the name in the complement position of the main clause, or another referent. Although most previous studies considered only the first two options, based on the results presented by Sorace and Filiaci (2006), it was decided to add this third alternative. The following is an example of a question:

(13) *Mateus harassed George when he interviewed him.*

*Who interviewed?*

- a) *Matthew*
- b) *George*
- c) *Another person*

To identify possible influences of the semantic relation between the verbs on the resolution of ambiguity, half of the examples were constructed with pairs of verbs considered “neutral” with respect to anaphoric interpretation, and the other half, in the absence of a better term, was considered as “biased”, i.e., pairs in which the semantic relation between the verbs seems to interfere with the interpretation of anaphora. Thus, in (13), it seems equally possible that the subject or the complement of the verb *to harass* may be the subject of the verb *to interview*. This sentence, therefore, was considered neutral. The following examples of questions, on the other hand, are considered biased:

(13) *Julia disrespected Paula when she betrayed her.*

*Who betrayed?*

- a) *Julia*
- b) *Paula*
- c) *Another person*

(14) *Laura bribed Lucy when she fined her.*

*Who fined?*

- a) *Laura*
- b) *Lucy*
- c) *Another person*

In (14), it seems more likely that the subject of the verb *to betray* is interpreted as corresponding to the subject of the verb *to disrespect*, that is, *Julia*. In (15), it seems more likely that the subject of the verb *to fine* is interpreted as coreferential to the complement of the verb *to bribe*, that is, *Lucy*. Instead of trying to avoid such possible influences, an attempt was made to analyse whether the semantic relation between the verbs could influence the resolution of ambiguity in the three languages of this study, either as L1 or L2.

The questionnaire was applied online in November 2019. The researcher visited the classrooms to ask students to participate in the research, and each professor sent the link by email to their students. Participation in the questionnaire was voluntary and anonymous. The socio-demographic data requested were the following: gender, age, variety of the first language (L1), number of years spent in Portugal, language of study at university (L2) and self-assessment of the level of proficiency in reading comprehension in the second language.

Except for English native speakers, who only replied to the L1 questionnaire, all other participants answered the L1 questionnaire, plus a brief reading proficiency test and the L2 questionnaire. The questions of the reading proficiency test were in accordance with the Common European Framework of Reference for Languages (CEFR) level descriptors (COUNCIL OF EUROPE, 2018) specified for Portuguese in the *Referencial Camões*<sup>2</sup>, for Spanish in the *Plan Curricular del Instituto Cervantes*<sup>3</sup> and for English in the *English Profile*<sup>4</sup>. In these questions, there were 20 sentences in each language (distributed evenly over levels A2, B1, B2 and C1) where some element was missing. This element was generally the verb or preposition required by the verb. For each question there were four options and, to avoid random answers, the alternative *I do not know*.

Although participation in the survey was anonymous, at the end of the questionnaire, anyone who wished to do so could write the name and email address for contact in order to receive information about the results of the study. Of the 23 participants, 16 asked to be sent the results of the survey.

## Results and discussion

Due to the small number of participants (N=23), who were divided between the three study languages (Portuguese, English, and Spanish), no correlation tests were possible. We have decided to carry out only descriptive statistical analysis in the SPSS v.26 software (IBM, 2019). Below, we present the analysis of the results.

In the foreign language proficiency test, both the seven Spanish learners and the five Portuguese learners hit more than 75% of the test, i.e. more than 15 questions. The average of Portuguese learners was 17.1, with a standard deviation of 1.3, and the average of Spanish learners was 18.1 questions, with a standard deviation of 1. English learners, despite being in the same semester of the course, showed more variation in the results, with a lower average of 15 and a higher standard deviation of 3.4. As there were 10 English learners, we decided to divide their answers into two groups: an intermediate level group with five students, who hit between 9 and 15 questions on the proficiency

---

<sup>2</sup> Available at: [https://www.instituto-camoes.pt/images/REFERENCIAL\\_ebook.pdf](https://www.instituto-camoes.pt/images/REFERENCIAL_ebook.pdf). Access on: 19 Oct. 2021.

<sup>3</sup> Available at: [https://cvc.cervantes.es/ensenanza/biblioteca\\_ele/plan\\_curricular/](https://cvc.cervantes.es/ensenanza/biblioteca_ele/plan_curricular/). Access on: 19 Oct. 2021.

<sup>4</sup> Available at: <https://www.englishprofile.org/english-grammar-profile/egp-online>. Access on: 19 Oct. 2021.

test, with a mean of 12.2 questions and a standard deviation of 2.3, and an advanced level group with five other students, who hit more than 15 questions, with a mean of 17.8 and a standard deviation of 1.3, as shown in Table 3 below.

**Table 3** – Proficiency tests

|                      | N  | Mean | SD  |
|----------------------|----|------|-----|
| Portuguese Total     | 5  | 17.1 | 1.3 |
| Spanish Total        | 7  | 18.1 | 1.0 |
| Advanced English     | 5  | 17.8 | 1.3 |
| Intermediate English | 5  | 12.2 | 2.3 |
| English Total        | 10 | 15.0 | 3.4 |

**Source:** Authors' elaboration.

For both the proficiency and the ambiguity tests, group averages were calculated and, since the N of the sample is less than 50 (MYERS; WELL; LORCH, 2013), the Shapiro-Wilk normality test (SHAPIRO; WILK, 1965) was carried out and the distribution of the results was concluded to be normal since  $p > 0.05$ .

We present below the tables with the percentages of responses of native speakers and learners of each language for the resolution of anaphora in the following situations:

- (i) anaphora (the anaphor comes after the antecedent) and cataphora (the anaphor comes before the “antecedent”);
- (ii) in which the anaphora is carried out using a pronoun (unstressed) or by means of an ellipsis (zero anaphora), in Portuguese and Spanish; in the case of English, using either a stressed or an unstressed pronoun;
- (iii) with the anaphor in the subordinate clause corresponding to the subject of the main clause, or to its complement, or to another referent;
- (iv) in “neutral” or “biased” sentences, i.e. where it is assumed that the lexical choice of verbs is decisive for the task.

To highlight in the tables the Position of Antecedent Hypothesis, the results that relate zero anaphora (in English, pronominal anaphora without emphasis) to the subject of the main clauses, and pronominal anaphora (pronominal anaphora with emphasis in English) to the complement of the main clauses will appear in bold. When no participant has identified the subject of the adverbial subordinate clauses as coreferential to an entity other than the subject or the complement of the main clauses, this alternative will be omitted from the tables.

We will start by presenting and discussing the results obtained on the resolution of ambiguous anaphora in Portuguese.

**Table 4** – Neutral sentences in Portuguese (%)

|           |            | Natives (N=13) |             | Learners (N=5) |             |
|-----------|------------|----------------|-------------|----------------|-------------|
|           |            | Subject        | Complement  | Subject        | Complement  |
| Anaphora  | Pronominal | 19.0           | <b>81.0</b> | 30.0           | <b>70.0</b> |
|           | Zero       | <b>65.4</b>    | 34.6        | <b>70.0</b>    | 30.0        |
| Cataphora | Pronominal | 34.6           | <b>65.4</b> | 50.0           | <b>50.0</b> |
|           | Zero       | <b>84.6</b>    | 15.4        | <b>50.0</b>    | 50.0        |

**Source:** Authors' elaboration.

As it can be seen in Table 4, most native speakers have interpreted the pronoun as coreferential to the complement of the main clauses and the ellipsis as coreferential to the subject of the main clauses, thus appearing to follow the Position of Antecedent Strategy. The learners, whose first language is Spanish, although presenting a similar behaviour to native speakers in solving the ambiguity in anaphoric contexts, in cataphoric contexts they did not show preference in the resolution strategy, probably due to the greater difficulty in processing cataphora.

Reversing the order of the clauses does not seem to influence the results of Portuguese native speakers, but it seems to cause more doubt in the identification of the antecedent by learners in neutral sentences. In biased sentences, however, such observation is not applicable.

In order not to extend the test too much, it was decided to create only 8 biased sentences, i.e. sentences in which the semantic interaction of the pair of verbs involved seemed to determine a preferential interpretation for the anaphora. Of the 8 sentences, half (2 sentences with zero anaphora and 2 with pronominal cataphora) indicated a preference in the interpretation of the subject of the subordinate clause as coreferential to the subject of the main clause, while the other half (2 sentences with pronominal anaphora and 2 with zero cataphora) indicated a preference in the interpretation of the subject of the subordinate clause as coreferential to the complement of the main clause. The expected preferences are marked in the table with an asterisk (\*).

**Table 5** – Biased sentences in Portuguese (%)

|           |            | Natives (N=13) |             |       | Learners (N=5) |              |
|-----------|------------|----------------|-------------|-------|----------------|--------------|
|           |            | Subject        | Complement  | Other | Subject        | Complement   |
| Anaphora  | Pronominal | 0              | <b>100*</b> | 0     | 20.0           | <b>80.0*</b> |
|           | Zero       | <b>84.6*</b>   | 15.4        | 0     | <b>90.0*</b>   | 10.0         |
| Cataphora | Pronominal | 69.0*          | <b>27.0</b> | 4.0   | 100*           | <b>0</b>     |
|           | Zero       | <b>50.0</b>    | 50.0*       | 0     | <b>10.0</b>    | 90.0*        |

**Source:** Authors' elaboration.

When comparing the data in Table 4 with those in Table 5, it can be seen that the semantic relation between the verbs influenced the resolution of ambiguity in Portuguese as L1 and as L2. As mentioned above, the biased sentences with pronominal anaphora and zero cataphora suggested that the antecedent should be the complement of the main clause, while the biased sentences with pronominal cataphora and zero anaphora suggested that the antecedent should be the subject of the main clause.

While, in neutral sentences, the learners were divided in the cataphora resolution, in biased sentences they were influenced by the semantic relation between the verbs (with the hypothesis marked with an asterisk), revealing the Position of Antecedent Strategy (with the hypothesis marked with bold). It is noticeable that native speakers, however, showed more resistance than language learners, especially for the interpretation of zero cataphora as coreferential to the complement of the main clause. Although the biased sentences influence native speakers' answers, they still seem to be influenced by the Position of Antecedent Strategy.

Now, let us see the results in Spanish.

**Table 6** – Neutral sentences in Spanish (%)

|           |            | Natives (N=5) |             |       | Learners (N=7) |             |       |
|-----------|------------|---------------|-------------|-------|----------------|-------------|-------|
|           |            | Subject       | Complement  | Other | Subject        | Complement  | Other |
| Anaphora  | Pronominal | 30.0          | <b>70.0</b> | 0     | 14.3           | <b>85.7</b> | 0     |
|           | Zero       | <b>80.0</b>   | 20.0        | 0     | <b>64.3</b>    | 35.7        | 0     |
| Cataphora | Pronominal | 50.0          | <b>40.0</b> | 10.0  | 14.3           | <b>78.6</b> | 7.1   |
|           | Zero       | <b>80.0</b>   | 20.0        | 0     | <b>50.0</b>    | 50.0        | 0     |

**Source:** Authors' elaboration.

As it can be seen in Table 6, in Spanish, native speakers tended to interpret zero anaphora and zero cataphora as coreferential to the subject of the main clauses, which was less evident for learners whose first language is Portuguese. Pronominal anaphora tended to be interpreted as coreferential to the complement of the main clause, but there did not seem to be a preferred interpretation for pronominal cataphora by native speakers.

In the neutral sentences in Spanish, speakers have not shown the same preference in all contexts. Although participants seem to follow the Position of Antecedent Strategy in most situations, natives were uncertain about pronominal cataphora, and learners were uncertain about zero cataphora. The results show a great difficulty in cataphora resolution, as it was already mentioned, since it was the only context in which the alternative *other* was also chosen.

**Table 7** – Biased sentences in Spanish (%)

|           |            | Natives (N=5) |             | Learners (N=7) |              |
|-----------|------------|---------------|-------------|----------------|--------------|
|           |            | Subject       | Complement  | Subject        | Complement   |
| Anaphora  | Pronominal | 0             | <b>100*</b> | 7.1            | <b>92.9*</b> |
|           | Zero       | <b>100*</b>   | 0           | <b>85.7*</b>   | 14.3         |
| Cataphora | Pronominal | 100*          | <b>0</b>    | 71.4*          | <b>28.6</b>  |
|           | Zero       | <b>20.0</b>   | 80.0*       | <b>85.7</b>    | 14.3*        |

**Source:** Authors' elaboration.

In relation to the biased sentences, both in Portuguese and Spanish, Spanish native speakers were more influenced by the semantic relation between the verbs than Portuguese native speakers. It can even be seen, from the data in Table 7, that learners interpreted cataphora zero in the opposite way to native speakers. While natives gave more importance to semantics for the choice of the antecedent, learners gave more importance to the syntax in the sentences with zero cataphora.

Comparing the results in both languages, in the biased sentences with zero cataphora, which suggested an interpretation of the antecedent as coreferential to the complement of the main clause, Portuguese students marked this option in 50% of the cases in Portuguese and in 14.3% of the cases in Spanish, while Spanish students marked this option in 90% of the cases in Portuguese and in 80% of the cases in Spanish.

The cataphoric neutral sentences were also related to a greater difficulty from the learners in identifying the antecedent. In Portuguese, learners were equally divided between the interpretation of the antecedent of the cataphora as the subject or complement of the main clause. In Spanish, while native speakers were divided on the interpretation of the antecedent of pronominal cataphora, learners were equally divided on the interpretation of the antecedent of zero cataphora as the subject or complement of the main clause.

It is clear from the above tables that both native speakers and learners prefer to relate zero anaphora to the subject of the main clause and for pronominal anaphora to be related to the complement of the main clause. In relation to cataphora, however, the preference is not so evident. This is also justified by the influence of biased sentences, which, in cataphoric contexts, suggested an interpretation contrary to the Position of Antecedent Strategy.

The verb combination, the antecedent and the order of the clauses seems to influence the resolution of ambiguity in Portuguese and Spanish. For English native speakers, however, the identification of the antecedent does not seem to be influenced by such variables.

**Table 8** – Neutral sentences in English (%)

|           |                  | Natives (N=5) |             | Advanced (N=5) |             | Intermediate (N=5) |             |
|-----------|------------------|---------------|-------------|----------------|-------------|--------------------|-------------|
|           |                  | Subject       | Complement  | Subject        | Complement  | Subject            | Complement  |
| Anaphora  | With emphasis    | 100           | <b>0</b>    | 40.0           | <b>60.0</b> | 60.0               | <b>40.0</b> |
|           | Without emphasis | <b>90.0</b>   | 10.0        | <b>70.0</b>    | 30.0        | <b>60.0</b>        | 40.0        |
| Cataphora | With emphasis    | 90.0          | <b>10.0</b> | 40.0           | <b>60.0</b> | 50.0               | <b>50.0</b> |
|           | Without emphasis | <b>100</b>    | 0           | <b>80.0</b>    | 20.0        | <b>60.0</b>        | 40.0        |

**Source:** Authors' elaboration.

**Table 9** – Biased sentences in English (%)

|           |                  | Natives (N=5) |              | Advanced (N=5) |              | Intermediate (N=5) |              |
|-----------|------------------|---------------|--------------|----------------|--------------|--------------------|--------------|
|           |                  | Subject       | Complement   | Subject        | Complement   | Subject            | Complement   |
| Anaphora  | With emphasis    | 60.0          | <b>40.0*</b> | 10.0           | <b>90.0*</b> | 20.0               | <b>80.0*</b> |
|           | Without emphasis | <b>100*</b>   | 0            | <b>70.0*</b>   | 30.0         | <b>60.0*</b>       | 40.0         |
| Cataphora | With emphasis    | 100*          | <b>0</b>     | 100*           | <b>0</b>     | 60.0*              | <b>40.0</b>  |
|           | Without emphasis | <b>60.0</b>   | 40.0*        | <b>90.0</b>    | 10.0*        | <b>50.0</b>        | 50.0*        |

**Source:** Authors' elaboration.

As it can be seen in Tables 8 and 9, there is no significant difference in the resolution of ambiguity depending on the type of antecedent or the order of the clauses in English. Although biased sentences influenced anaphora resolution by native speakers when the biased interpretation was directed to the complement of the main clause, most native speakers tend to interpret the pronoun as coreferential to the subject of the main clause, confirming the results of Valenzuela, Liceras, and Morelos (2011) had shown.

The significant variable when comparing results in English, however, unlike in Portuguese and Spanish, was the participants' level of proficiency in the language. It was not possible to identify any difference between proficiency levels in the Portuguese and Spanish groups, as non-native speakers had advanced proficiency level in reading comprehension in L2. For English non-native speakers, however, as there were two groups with different proficiency levels (one intermediate and one advanced), more differences were found when comparing the results of the intermediate group with the results of the advanced group and the native speakers.



A common result in the three languages was the general choice by the participants to relate the subject of the subordinate clauses only to the subject or the complement of the main clauses. In the study of Sorace and Filiaci (2006), the participants interpreted pronominal cataphora as corresponding to an entity not mentioned in the main clause. However, the authors presented images with three possible referents for the ambiguous sentences, and the illustration of the three referents may have influenced the resolution of ambiguity. As our study did not count with images, but only involved ambiguous sentences, participants did not have a preferential interpretation to relate the anaphora or the metaphor to a referent that did not present an antecedent in the sentence.

In English sentences, no participant chose an antecedent other than the subject or the complement of the main clauses. In Portuguese and Spanish, although not significant, the resolution of ambiguous anaphora as related to an entity that was not the subject or the complement of the main clause occurred only in relation to pronominal cataphora. In the neutral sentences in Spanish, this option was selected in both L1 and L2. Portuguese native speakers also selected this option in the biased sentences in Portuguese.

## **Final considerations**

This study investigated the strategies used in the resolution of ambiguous anaphora in Portuguese, English, and Spanish both as L1 and as L2. It was concluded that English speakers prefer to interpret the anaphoric pronoun in the subject position of the subordinate clause as coreferential to the subject of the main clause, regardless of whether the pronoun receives emphasis or not. Learners with intermediate proficiency level in English have distanced themselves from native speakers and from learners with advanced proficiency level.

It has been confirmed that in relation to pronominal anaphora and zero anaphora, Portuguese and Spanish speakers follow the Position of Antecedent Strategy, interpreting zero anaphora in the subject position of the subordinate clause as coreferential to the subject of the main clause; and interpreting pronominal anaphora as coreferential to the complement of the main clause. In neutral sentences with zero cataphora, there was a great difficulty on the part of Portuguese and Spanish learners to identify the antecedent.

Although, in general, participants chose only the subject or the complement of the main clauses as antecedents, the choice of another referent occurred only in sentences with pronominal cataphora in Portuguese and Spanish. The inversion of the order of the clauses, therefore, seems to make the resolution of ambiguity more difficult.

This analysis was based on a pilot study conducted in November 2019 with a small sample of 23 students from the University of Algarve, in Portugal. In 2020, the research will be carried out with a larger sample in both Brazil and Portugal to compare data obtained with participants from different levels of proficiency and from two varieties of Portuguese: Brazilian Portuguese and European Portuguese.

The carried out research contributes to the state of the art not only by explaining criteria for the construction of sentences and by carrying out a comparative study in three languages, but also by demonstrating with quantitative data that the semantic relation between the verbs influences the resolution of ambiguity, especially for Spanish native speakers. The data obtained in this study will later be compared with a larger sample data to check the consistency of the results.

## Acknowledgements

We would like to thank all the people who contributed to this study by answering or disseminating the survey, as well as the university where it was conducted, the University of Algarve. We would also like to thank Professor Elena Valenzuela, for making available material from her research, in which we were partly inspired by this work.

BRUSCATO, A.; BAPTISTA, J. Resolução de Ambiguidade Anafórica em Português, Inglês e Espanhol (estudo-piloto). *Alfa*, São Paulo, v.65, 2021.

- *RESUMO: Este artigo apresenta um estudo-piloto no qual se pretende analisar as diferenças na resolução de ambiguidade anafórica em português, inglês e espanhol como primeiras línguas (L1) e como segundas línguas (L2). Para a obtenção dos dados, foi elaborado um questionário, divulgado entre os estudantes nacionais e estrangeiros da Universidade do Algarve, Portugal. Conclui-se que os falantes nativos da língua inglesa costumam interpretar o pronome anafórico em posição de sujeito da oração subordinada como correferente ao sujeito da oração principal, independentemente de o pronome receber ou não ênfase. Já os falantes de português e espanhol parecem seguir a Estratégia da Posição do Antecedente (CARMINATI, 2002) nos casos de anáfora, mas a preferência torna-se menos evidente ao inverter a ordem das orações (casos de catáfora). Além de considerar variáveis sintáticas, as frases foram elaboradas de modo a investigar a influência das relações semânticas entre os verbos para a resolução de ambiguidade anafórica.*
- *PALAVRAS-CHAVE: Anáfora; espanhol; inglês; português; resolução de ambiguidade.*

## REFERENCES

AKMAJIAN, A.; JACKENDOFF, R. Coreferentiality and stress. *Linguistic inquiry*, Cambridge, v.1, n.1, p. 124-126, 1970.

ALONSO-OVALLE, L.; FERNÁNDEZ-SOLERA, S.; FRAZIER, L.; CLIFTON, C. Null vs. overt pronouns and the topic-focus articulation in Spanish. *Italian Journal of Linguistics*, Ospedaletto, v.14, n.2, p. 151-170, 2002.

BAPTISTA, J. ViPER: uma base de dados de construções léxico-sintáticas de verbos do Português Europeu. *In: ENCONTRO DA APL*, 28., 2013, Coimbra. **Actas [...]**, Coimbra: APL, 2013. p. 111-129.

CARMINATI, M. **The processing of Italian subject pronouns**. 2002. Thesis (Doctor of Philosophy) - University of Massachusetts, Amherst, 2002.

CHOMSKY, N. **Lectures on government and binding**. Dordrecht: Foris, 1981.

COUNCIL OF EUROPE. **Common European Framework of Reference for Languages: Learning, teaching, assessment: Companion volume with new descriptors**. Strasbourg: Council of Europe Publishing, 2018.

CRAWLEY, R.; STEVENSON, R.; KLEINMAN, D. The use of heuristic strategies in the interpretation of pronouns. **Journal of Psycholinguistic Research**, New York, v.19, n.4, p. 245-264, 1990.

HALLIDAY, M.; HASAN, R. **Cohesion in English**. London: Longman, 1976.

IBM. **SPSS for Windows**. Version 26, 2019.

JEGERSKI, J.; VANPATTEN, B.; KEATING, G. Cross-linguistic variation and the acquisition of pronominal reference in L2 Spanish. **Second Language Research**, London, v.27, n.4, p. 481-507, 2011.

LOBO, M.; SILVA, C.. Ambiguidade pronominal em orações adverbiais do Português Europeu. **Revista da Associação Portuguesa de Linguística**, [s.l.], v.2, n.2, p. 319-338, 2016.

LUJÁN, M. Stress and binding of pronouns. **Chicago Linguistic Society**, Chicago, v.22, n.2, p. 248-262, 1986.

MORGADO, S. Qual a importância da informação semântica na resolução de sujeitos pronominais no processamento offline. *In: ENCONTRO DA APL*, 28., 2013, Coimbra. **Actas [...]**, Coimbra: APL, 2013. p. 493-507.

MYERS, J.; WELL, A.; LORCH R. **Research design and statistical analysis**. 3rd. ed. New York: Routledge, 2010.

NASCENTES, A. **Dicionário etimológico da língua portuguesa**. Rio de Janeiro: Oficinas gráficas do Jornal do Comércio, 1955. 2 v.

RIZZI, L. **Issues in Italian Syntax**. Dordrecht: Foris, 1982.

SHAPIRO, S.; WILK, M. An analysis of variance test for normality. **Biometrika**, Oxford, v.52, n.3, p. 591-611, 1965.

SORACE, A.; FILIACI, F. Anaphora resolution in near-native speakers of Italian. **Second Language Research**, London, v.22, n.3, p. 339-368, 2006.

SZWEDEK, A. The role of sentence stress in the interpretation of coreferentiality in English and Polish. *In*: FISIĄK, J. (ed.). **Theoretical Issues in Contrastive Linguistics**. Amsterdam: John Benjamins Publishing, 1980. p. 377-387.

TORIBIO, A. Setting parametric limits on dialectal variation in Spanish. **Lingua**, Amsterdam, n.110, p. 315-341, 2000.

TSIMPLI, I.; SORACE, A.; HEYCOCK, C.; FILIACI, F. First language attrition and syntactic subjects: a study of Greek and Italian near-native speakers of English. **International Journal of Bilingualism**, London, v.8, n.3, p. 257-277, 2004.

VALENZUELA, E.; LICERAS, J.; MORELOS, L. Ambiguous anaphora in L2 English and L2 Spanish. *In*: CONFERENCE ON LANGUAGE DEVELOPMENT, 36., 2011, Boston. **Annals** [...], Boston, 2011.

## Appendix

### Sentences in Portuguese

1. Ana insultou Carolina quando ela a chantageou.
2. Quando a chantageou, Ana insultou Carolina.
3. Maria abraçou Joana quando a contratou.
4. Quando ela a contratou, Maria abraçou Joana.
5. António estrangulou Carlos quando o torturou.
6. Quando ele o torturou, António estrangulou Carlos.
7. Mateus assediou Jorge quando ele o entrevistou.
8. Quando o entrevistou, Mateus assediou Jorge.
9. Júlia desrespeitou Paula quando a traiu.
10. Quando ela a traiu, Júlia desrespeitou Paula.
11. Laura subornou Lúcia quando ela a multou.
12. Quando a multou, Laura subornou Lúcia.
13. João algemou Pedro quando o sequestrou.
14. Quando ele o sequestrou, João algemou Pedro.
15. Lucas castigou Luís quando ele o confrontou.
16. Quando o confrontou, Lucas castigou Luís.
17. Artur ajudou Alice quando ela o chamou.
18. Quando o chamou, Artur ajudou Alice.
19. Diana beijou Daniel quando o visitou.
20. Quando ela o visitou, Diana beijou Daniel.

### Sentences in Spanish

1. Ana insultó a Carolina cuando ella la chantajeó.
2. Cuando la chantajeó, Ana insultó a Carolina.
3. María abrazó a Juana cuando la contrató.
4. Cuando ella la contrató, María abrazó a Juana.
5. Antonio estranguló a Carlos cuando lo torturó.
6. Cuando él lo torturó, Antonio estranguló a Carlos.
7. Mateus acosó a Jorge cuando él lo entrevistó.
8. Cuando lo entrevistó, Mateus acosó a Jorge.
9. Julia irrespetó a Paula cuando la traicionó.
10. Cuando ella la traicionó, Julia irrespetó a Paula.
11. Laura sobornó a Lucía cuando ella la multó.
12. Cuando la multó, Laura sobornó a Lucía.
13. Juan esposó a Pedro cuando lo secuestró.
14. Cuando él lo secuestró, Juan esposó a Pedro.
15. Lucas castigó a Luís cuando él lo confrontó.

16. Cuando lo confrontó, Lucas castigó a Luís.
17. Arturo ayudó a Alice cuando ella lo llamó.
18. Cuando lo llamó, Arturo ayudó a Alice.
19. Diana besó a Daniel cuando lo visitó.
20. Cuando ella lo visitó, Diana besó a Daniel.

### Sentences in English

1. Anna insulted Carolina when SHE blackmailed her.
2. When she blackmailed her, Anna insulted Carolina.
3. Mary hugged Joanna when she hired her.
4. When SHE hired her, Mary hugged Joanna.
5. Anthony strangled Carl when he tortured him.
6. When HE tortured him, Anthony strangled Carl.
7. Matthew harassed George when HE interviewed him.
8. When he interviewed him, Matthew harassed George.
9. Julia disrespected Paula when she betrayed her.
10. When SHE betrayed her, Julia disrespected Paula.
11. Laura bribed Lucy when SHE fined her.
12. When she fined her, Laura bribed Lucy.
13. John handcuffed Peter when he kidnapped him.
14. When HE kidnapped him, John handcuffed Peter.
15. Luke castigated Louis when HE confronted him.
16. When he confronted him, Luke castigated Louis.
17. Arthur helped Alice when SHE called him.
18. When she called him, Arthur helped Alice.
19. Diana kissed Daniel when she visited him.
20. When SHE visited him, Diana kissed Daniel.

Received on April 29, 2020

Approved on October 20, 2020