Revista de Estudos da Linguagem, Belo Horizonte, v.25, n.3, p. 1647-1684, 2017

The processing of the reduplicated numeral sohoji-sohoji in Karajá: An ERP investigation during sentence comprehension

O processamento do numeral reduplicado sohoji-sohoji em Karajá: uma averiguação de ERP durante a compreensão de sentenças[1]

Cristiane Oliveira da Silva
Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
cristianeolivers@gmail.com

Resumo: Este trabalho tem como objetivo averiguar o processamento do numeral distributivo *sohoji-sohoji*, formado pela reduplicação do numeral *um* em Karajá, sob uma perspectiva transdisciplinar aproximando os campos da linguística teórica e da neurobiologia. O experimento de neurociência da linguagem contribui para elucidar a complexidade do processamento de sentenças com *sohoji-sohoji* quando pareado com diferentes cenários quantitativos. Utiliza-se a técnica de extração de ERP (potenciais bioelétricos relacionados a eventos) para observação e extração dos dados de latência visando a análise estatística dos dados cronométricos monitorados *online* por um eletroencefalograma (EEG). Os resultados das análises de variância e teste *t* demonstraram que a leitura evento-distributiva é a *default* porque a violação de exaustividade nos cenários testados apresentou maior complexidade de processamento do que a violação de exaustividade de participantes. Pluralidade, cardinalidade e diferenciação de eventos também apresentaram efeitos significativos durante a compreensão de sentenças com *sohoji-sohoji*.

Palavras chave: Karajá; distributividade; ERP; processamento quantificacional.

[1] Work based on the Doctoral Dissertation defended at UFRJ in 2016; Advisor: Marcus Maia and Coadvisor: Aniela Improta França.

eISSN: 2237-2083
DOI: 10.17851/2237-2083.25.4.1647-1684

Abstract: The aim of this paper is to investigate the processing of the reduplicated numeral *sohoji-sohoji* formed by the reduplication of the numeral *one* in Karajá, within an interdisciplinary approach, bringing together formal linguistics and neurobiology. The neuroscience of the conducted language experiment sheds new light on the quantification processing complexity of *sohoji-sohoji* when different quantitative scenarios were previously presented. We use the Event-Related Brain Potential (ERP) extraction technique to observe and extract the wave's amplitude in order to conduct a statistical analysis from the chronometric data monitored online by the electroencephalogram (EEG). The results from statistical analyses (ANOVA and *t* test) conducted in this study suggest that the event-distributive interpretation is the first to be activated, because the violation of event item exhaustiveness was more complex than was the violation of subject exhaustiveness. Plurality, cardinality, and event differentiation also displayed significant effects during Karajá distributive sentence comprehension.

Keywords: Karajá; distributive property; ERP; quantification processing.

Received on January 12, 2017

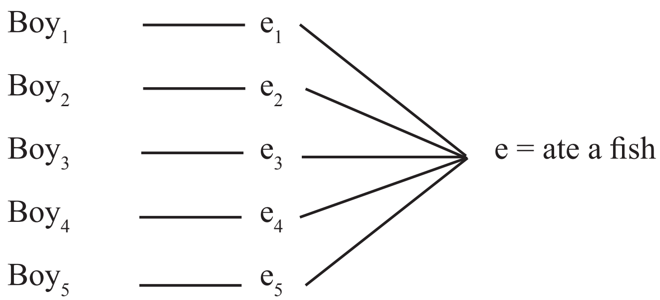
Approved on May 8, 2017

1. Introduction

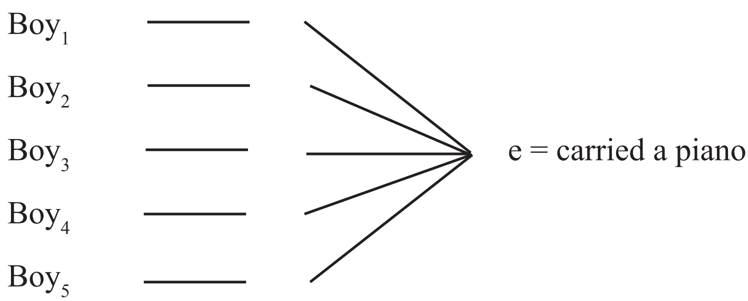
The quantification of events or individuals is a common operation in natural languages. However, each linguistic system has distinct mechanisms for expressing notions related to quantity. The study of quantifiers in natural languages is a topic of interest in different areas of knowledge, as philosophy and logic, for example, having a range of classical works that attempt to ascertain the functioning of linguistic elements to express quantity. In this paper, we present a processing study that used the Event-Related Brain Potentials (ERP) technique to investigate neurobiological responses during the processing of distributive sentences with the *sohoji-sohoji* quantifier, a reduplicated numeral in the Karajá language.

Distributive quantifiers are responsible for a mapping operation between the subevents and individuals of a statement. Therefore, a prototypical distributive reading should operate on individuals and subevents under the condition of differentiation between mapped individuals and subevents. As shown in example (1), for each individual x such that x is a *boy*, there will be a different sub-event of *eat a fish*.

(1) Each boy ate a fish.



(2) The boys carried a piano.



However, delimiting certain types of events does not seem to be a trivial task and non-differentiation of the subevents could then result in a collective reading, as in example (2). Tunstall (1998) also emphasizes the existence of a gradient between distributive and collective readings, i.e., there would be a continuous scale between prototypical distributive and collective readings in which intermediate readings would be available, whether partially distributive or partially collective. Kratzer (2003, 2007) treats these intermediate readings as cumulative.

Krifka (1998) and Kratzer (2003, 2007) argue in favor of the universal cumulative property of lexical items, although this does not guarantee that the lexicon is cumulative in the language syntax. In this light, assuming that lexical items are also cumulative in Karajá syntax, a cumulative reading would always be available in that language. Hence, we could speculate that the processing of a cumulative reading would present a lower processing cost, except when there is an element that delimits the quantifiable scope of the sentence, as, for example, a distributive operator. Therefore, by using the *sohoji-sohoji* reduplicated numeral, a distributive reading for the statement would be forced, even if other readings could be accommodated. In this bias, two questions can be raised in relation to the Karajá language: (i) Would it be possible to process intermediate readings in the presence of the reduplicated numeral? (ii) Would pairing intermediate readings of sentences with *sohoji-sohoji* quantifier present a higher processing cost? We confirmed in a pilot study based on scenario descriptions that scenarios containing intermediate readings could be described by the participants using the *reduplicated* numeral. This fact led to the development of an experiment that would make it possible to verify whether there would be sentence processing differences between the different readings available when paired with the reduplicated numeral in Karajá.

Thus, an ERP experiment was performed to investigate the processing of the *sohoji-sohoji*. The experiment was based on a picture matching design. We attempted to assess the course of sentence processing presented by capturing the reaction times and the neurobiological responses during the experiment. To this end, we ascertained the processing of sentences containing the reduplicated numeral after showing the distributive, collective, and partially distributive scenarios. As shown below, the properties of cumulativity, distributivity and cardinality exert effects during the processing and interpretation of sentences with *sohoji-sohoji* in Karajá. Such evidence seems to run in line with current theories about quantification processing and the semantics of distributional numerals, which have been observed in other natural languages.

2. The reduplicated numeral in Karajá (and other languages)

The Karajá language (Karajá family, Macro-Je stock) presents numeral reduplication to express the distributive property, as can be seen in example (2) below. Numeral reduplication brings to light two types of distributive readings, according to the interpretations presented in (3).

(2)	*weryry*	*sohoji-sohoji*	*bola-di*	*r-e-hu-ra*
	BOY	ONE-ONE	BALL-POSP	3A-VT-TO.THROW-PST.REC

Lit: 'Boy threw the ball one by one

(3)	a. 'The boys threw a ball one at a time'	event-distributive
	b. 'Each boy threw a ball	participant-distributive

While the event-distributive reading (3a) implies that each individual has thrown the ball in distinct periods of time, the participant-distributive reading (3b) does not present this requirement for the sentence interpretation. In turn, a collective reading does not seem to be allowed, i.e., a reading in which all the boys have thrown a single ball at the same time.

The difference between the readings above seems to come from the semantic ambiguity triggered by the reduplicated numeral, *sohoji-sohoji*. One of the objectives of this paper is to determine whether there would be a preference for some of the readings presented in (3) and whether a collective reading would be immediately discarded, given its incompatibility with the semantics of the distributive numeral, during the processing of distributive sentences in Karajá,.

The possibility of a participant-distributive reading requires that the distributive operate over individuals. The impossibility of combining the reduplicated numeral with mass nouns in Karajá, which do not allow the individuation of entities because it is a substance, can be confirmed by impossibility of exemple (4) below. Sentence (4) demonstrates that the substance *bèè* (water) is incompatible with the distributive operation carried out by *sohoji-sohoji*.

(4)	**Tii*	*bèè*	*sohoji-sohoji-my*	*r-e-wy-reri.*
	3P	water	one-one-POSP	3P-VT-to.transport-PRS.CONT

He carries the water one at a time/ He carries each water

Numeral reduplication as a distributive operation strategy is relatively common in several languages. According to the WALS[2] mapping, 85 from 251 languages investigated, present a numeral reduplication strategy to denote adnominal distributive property. Gil (1982) also shows that all languages have some resource to express adverbial distributive property by numeral morphology. It should be noted, however, that these studies cover very few indigenous languages and the Karajá language was not investigated by these papers.

Karitiana (Arikém family) is another indigenous language that has reduplicated numerals to express distributive property. As in Karajá, reduplicated numerals in Karitiana generate different possibilities of interpretation for the same sentence as exemplified below in (5). Muller (2012) assumes, based on syntactic tests, that distributive numerals in Karitiana are adverbial distributive operators. For the author, the two available readings are the result of different modes of individuation that is possible for the subevents in the verbal predicate denotation.

(5)	Õwã	*nakakot*	*sypomp.sypomp*	*opokakosypi*
	BOY	BROKE	TWO.TWO	EGG

‘Each boy broke two eggs’ – participant-distributive

‘The boys broke two eggs at a time’ – event-distributive

Muller (2012, p. 235)

However, in analyzing the numeral morphology in Tlingit (Na-Dene family – Canada), Cable (2014) proposes that numeral distribution always allows participant-distributive and event-distributive scenarios to be licensed, similar to what is found in Karajá and Karitiana.

> Amongst other facts, I show here that the syntactic attachment site of the distributive numeral has no effect upon the kind of distributive scenario the sentence describes. That is, no matter whether the distributive numeral is adverbial or adnominal, the sentence may be true in either participant-distributive or event-distributive scenarios. (CABLE, 2014, p. 8)

Cable (2014) proposes two semantic denotations, based on a scenario description test, a distributive-adverbial and a distributive-adnominal, for numeral morphology in Tlingit. However, the author argues that both denotations can generate distributive readings about events as well as individuals. Therefore, it is concluded that the ambiguity generated by distributive numerals is recurrent in the languages that have this morphological strategy to express the distributive property in their statements. Based on the data, the ambiguity of the distributive numerals does not seem to be the result of syntactic aspects, but instead the effect of semantic operations which are peculiar to these elements.

Thus, with the intention to ascertain possible differences in processing between possible (and impossible) readings for the reduplicated numeral in Karajá, we conducted an EEG experiment, as presented in Section 4 below. The results can help explain the cognitive mechanisms during numeral distributive processing. The processing of sentences that present quantifying elements is known to be a very complex procedure, since it requires interface between different cognitive modules, such as the syntactic and the quantificational. Therefore, it is part of cognitive research and neuroscience agenda to distinguish the neural mechanisms underlying access to different aspects of semantic memory during sentence comprehension (JIANG *et al*., 2008).

3. Theoretical assumptions

All natural languages known to date have mechanisms for quantifying sentences. As stated in the introduction, one of the objectives of this paper is to understand the nature and processing of one of the quantifying elements in Karajá, the reduplicated numeral *sohoji-sohoji*, in order to contribute to understanding how quantification works in natural languages and what are the basic principles governing these elements within human languages. The ability to quantify linguistic elements is commonly expressed through operators called quantifiers, which, according to the definition given by Lyons (1977, p. 455), are modifiers that, when combined with names, produce expressions whose reference is determined by means of the size of the set of elements, or in terms of the amount of the substance being referred to. However, it is known that there are also different types of non-nominal quantifiers, as is the case for certain quantity-expressing adverbs.

From a relational perspective, quantifiers denote binary relations between sets. Quantifiers are entrusted with the task of relating two sets of distinct elements, in this case, the set of individuals and the set of properties. In terms of Formal Semantics, Heim (1982) states that a quantifier, when relating two sets of elements, forms a triad in which it occupies the position of operator of the quantified structure, as shown in (6) below.

(6) a. Each monkey ate a banana.

b. **for** Each$_x$ **such that** [*x* is a monkey (∃$_y$) **such that** *y* is a banana **and** *x* ate *y*

operator	restrictive clause	nuclear scope

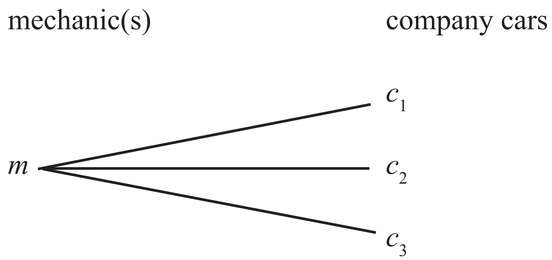
"For each x such that x is a monkey there is a y such that y is a banana and x ate y".

Therefore, from a formal viewpoint, it seems appropriate to establish that quantifiers are operators that perform the action of relating two sets belonging to seemingly distinct classes. In the case of example (6), the quantifier *each* relates the involvement of the individuals, which belong to the monkey set, in subevents of "eating a banana". In this area of linguistics studies, analyses of scope ambiguities, caused mainly by the interaction between quantified nominal phrases, have flourished.

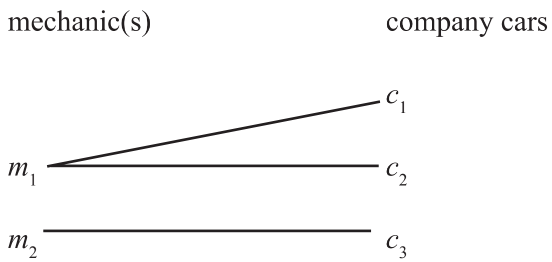
(7) A mechanic repaired every car of the company.

The sentence (7) above would generate at least two possible interpretations: (7a) there is a (single) mechanic who repaired every car in the company and (7b) each company car was repaired by a mechanic (any). These two interpretations can be schematized as follows:

(7a)



(7b)



Syntactically, each interpretation above is represented by a different Logical Form structure that will be generated from the Nominal Phrase raising that is in the verbal complement position. Therefore, for the interpretation of sentences with quantified NPs in the object position, an abstract syntactic structure is necessary to serve as input for the semantic interpretation of sentences with this type of structural component (MAY, 1977, 1985).

Distributive property is an operation that occurs mainly when the nucleus of the Determiner Phrase is occupied by a distributive quantifier and combined with a predicate. The predicate is then understood as applied to each individual member of the quantified set. Therefore, the operation of a distributive quantifier is to map subevents to each available atom of speech. Obviously, this is a very simplistic definition that does not account for a series of distributive mappings that involves scope ambiguity in doubly quantified sentences, for instance. However, it seems to be a statement clear and simple enough to explain the relationship between DPs that contain distributive quantifiers and events to which they are applied, as in "each monkey ate a banana." In this sentence, "each" takes the individual members of the restrictor set {x: x is monkey}, with the predicate [ate banana] being applied to each individual member of {x: x is monkey}. Furthermore, a distributive quantifier ends up pluralizing the event of the sentence in which it occurs, since in a distributive structure, at least two subevents of the available set of events are required, i.e., it requires that the R relationship between individuals and events is greater than 1. Then, for sentence (6) above, the meaning can be interpreted as follows: there is an event *e* and for each individual member of {x: x is monkey} there is a sub-event *e'* which is part of *e* and such that in *e'm* "ate banana".

Distributive quantifiers such as *each* distribute over individual entities. Therefore, when applied to uncountable noun phrases, such as those occupied by mass names of the substance type, they generate a ungrammatical sentence, as illustrated in (8) below, since it would not be possible to individualize its parts, except under coercion or when a unit of measurement is attached to the noum in question. In Karajá, the distributive *sohoji-sohoji* exhibits the same pattern of behavior, already expected for a quantifier of this nature, given the grammaticality of (9a) and the ungrammaticality of (9b).

(8) *He brought each water.

(9a) *Tii hãdoroo sohoji-sohoji-my r-e-wy-reri.*

3P MURICIUM[2] UM-POSP 3P-VT-TO.CARRY-PRS:CONT

'He takes murici one by one'/ 'He takes each murici'

(9b) **Tii bèè sohoji-sohoji-my rewyreri.*

3P WATER UM-UM-POSP 3P-VT-TO.CARRY-PRS:CONT

'He takes water one by one' / 'He takes each water'

Distributive quantifiers also force a distributive reading of the event (subevents) of the sentence in which it is inserted. Therefore, when combined with collective predicates, they generate ungrammatical sentences (DOWTY, 1987). For example, sentence (10) below requires the predicate "be piled up in the corner of the room" be applied to a set of individuals {x: x books} in (10a) and (10b), and not to individual members of this set, given the ungrammaticality of (10c) and (10d).

[2] A Brazilian fruit from the Byrsonima crassifólia specie.

(10a) The books are piled up in the corner of the room.

(10b) All the books are piled up in the corner of the room.

(10c) *A book is piled up in the corner of the room.

(10d) *Each book is piled up in the corner of the room.

As expected in typically distributive predicates, distributive quantifiers can be applied normally, as in (11) below. Likewise, if a distributive quantifier is not applied to a sentence whose predicate is distributive, a predicate of this nature would still force a distributive reading, even if it does not contain a DP composed of a nucleus quantified by a distributive element in its structure, as seen in (12 and 13) below.

(11) Every girl slept.

(12) All the girls slept.

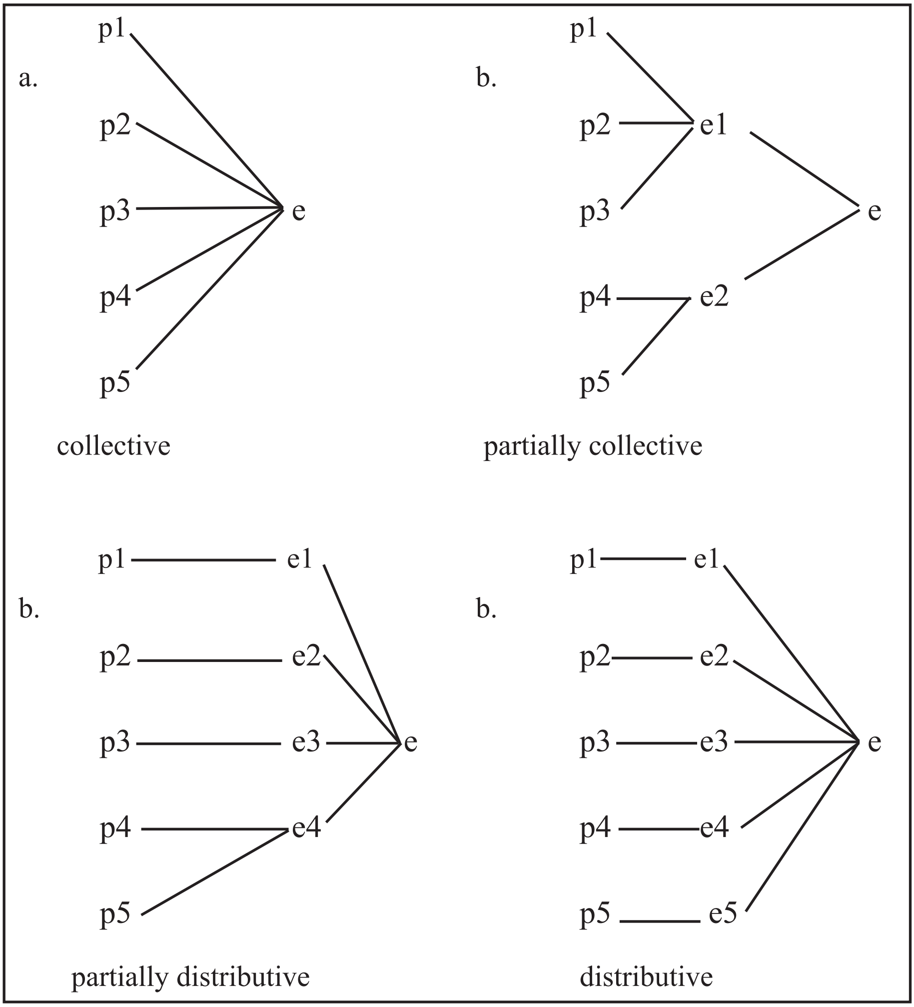
(13) The girls slept.

Thus, according to the above data, it seems that a distributive Reading must operate not only over individuals > 1, but also over subevents >1, thus creating the condition of the differentiation between individuals and subevents – though it is complicated to limit subevents in certain predicates. Non-differentiation of subevents results in a collective reading. However, there are not only collective readings and standard distributive readings. According to Tunstall (1998), distributive-collective event structures form a continuous scale in which partially collective and partially distributive readings are at the middle, while the completely collective and the completely distributive readings are at the ends. Kratzer (2003, 2007) treats intermediate readings of this scale as cumulative. Following Krifka (1998), Kratzer (2003, 2007) introduces the concept of cumulative property as a crucial point for the denotation of predicates in natural languages and argues in favor of the Universal Cumulative property of lexical items that can be conferred in the following section.

> "If there is a basic logical-conceptual predicate 'red', for example, that is true of my hat and your scarf (two singularities), then the Cumulativity Universal says that that very same predicate is also true of the sum of my hat and your scarf (a plurality)." (KRATZER, 2003, p. 12).

This means that, in simple predicates, a cumulative reading would be available without higher processing costs, as it is inherent and available to lexical items. However, the presence of a distributive operator would result in a preference for a distributive reading. Moreover, the cumulativeness inherent in lexical items would also explain the fact that a series of experiments presents evidence of a higher processing cost during the understanding of quantified predicates. This would occur because the presence of quantifying elements would reveal greater complexity for the initial semantic processing of sentences, which would be cumulative by default. The ERP experiment presented in the next section seems to point in this direction. Look at the the following diagrams, taken from Tunstall (1998, p. 97), which examine the possible intermediate readings between the standard collective and distributive readings. This scale allowed the author to demonstrate that the main difference between the semantics of *each* and *every* concerns the differentiation condition, which is stronger for *each* (every subevent should be distinct) than for *every* (at least two events must be distinct).

FIGURE 1 – Intermediate readings diagram by Tunstall (1998)



(14) Maria picked up all the folders.

Imagine the following scenario for a sentence like (14) above: there are five document folders on a table such that Maria picked up all these folders. In this scenario, there are a number of possible readings according to the scale shown above. Reading (a) – **collective** – predicts that Maria picked up all five folders at one time, i.e., there was only one event of "pick up the folders". Reading (b) – **partially collective** – suggests the folders were picked up in small groups, for example, Maria picked up three folders and then picked up two more folders, thereby generating two subevents of "pick up the folders", as shown in the above chart. Reading (c) – **partially distributive** – requires there be subevents applied to individual members, but concomitantly allows collective subevents, for example, Maria picked up three folders one at a time and then two folders at one time, i.e., four subevents of "pick up the folders". Finally, reading (d) – **completely distributive** – requires each folder to be picked up individually, thus generating five "pick up the folders" subevents.

In the following section, we will investigate whether scenarios that have a partial distributive property can be paired with sentences containing the reduplicated numeral *sohoji-sohoji* and what would be the cognitive cost of this procedure. In addition, we aim to demonstrate the acceptability (or otherwise) of different quantified scenarios when the participant was exposed to a sentence operated by *sohoji-sohoji*. The offline and online results corroborate the existence of a gradient of acceptability for different distributive readings, in which completely distributive and singular scenarios are at opposite ends of the scale. We will see that, although the behavioral test results elicited acceptability for the different scenarios, including those that are typically non-distributive, the measurements of time (ms) and amplitude (μV) seem to provide significant data for our hypothesis that, although different readings may be possible for the *sohoji-sohoji* quantifier, the standard distributive reading is the first one activated by distributive sentences in Karajá whose operator is *sohoji-sohoji*, i.e., the one that requires less cognitive cost during sentence processing.

4. The processing of the reduplicated numeral in Karajá

An experimental design was set up to investigate the processing of the reduplicated numeral, which made it possible to collect online and offline data. The task consisted of pairing different controlled scenarios, which were pictures presented on a computer screen, with distributive sentences recorded in audio. This method is commonly known as picture-matching design.

The objective of this experiment[4] was to observe whether standard distributive contexts would present processing differences in relation to other quantitative contexts (partially distributive, collective, non-exhaustive, extra-element, and uniqueness) during sentence comprehension with *sohoji-sohoji*. This is because we have observed in previously verified pilot tests that non-prototypically distributive scenarios were accepted by the participants as possible scenarios for sentences containing *sohoji-sohoji*. Therefore, the present experiment aimed to ascertain whether the acceptance of different quantitative scenarios for sentences with *sohoji-sohoji* would be the result of a reflexive process involving independent variables, such as interpretation and accommodation, subsequent to the parsing of sentences, or whether the acceptance of different scenarios allowed by the quantifier *sohoji-sohoji* would be related to the intrinsic semantic properties generated by this reduplicated numeral and, therefore, to be observed at a more reflexive level of sentence comprehension, called online processing.

4.1 Participants

Twenty-two subjects (16 men and 6 women), with corrected or normal vision, normal hearing, right-handed individuals (OLDFIELD, 1971), native speakers of the Karajá language, inhabitants of the Btõiry or Hawalò villages (Ilha do Bananal, TO) participated in this experiment. All participants stated that they had not consumed alcohol or taken any prescription or illicit drug 24 hours before the experiment and had completed elementary education. The mean age of the participants was 26 years.

4.2 Method and Materials

The experiment consisted of presenting the following sequence of events: a picture appeared on a CRT monitor, followed by four audios – corresponding to the one, two, three, and four sequences of a sentence in Karajá. All experimental sentences presented the subject-distributive-locative-verb structure respecting the typological pattern of the language, as in example (15) below. After listening to the audios, the volunteers' task was to decide whether or not the picture would be compatible with the sentence heard. This method is known in experimental psycholinguistics as picture-matching design.

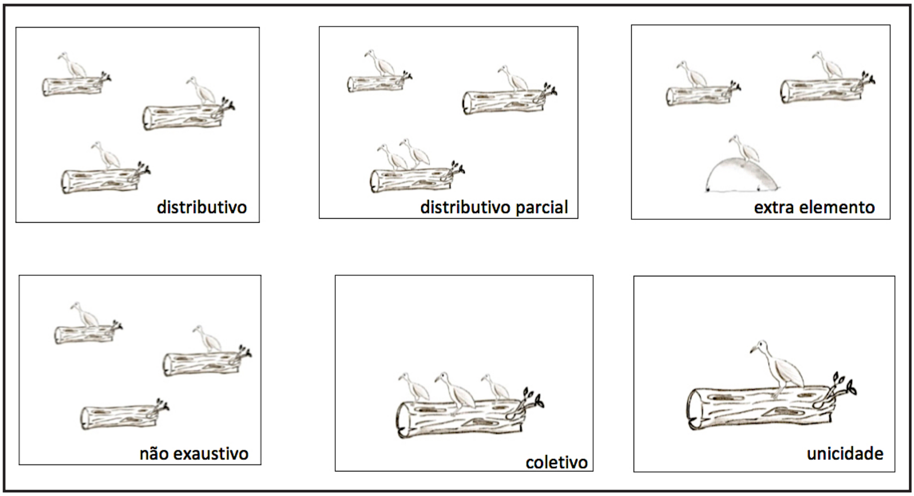
(15) *Krukru sohoji-sohoji òra-tyreki r-y-i-reri.*
PÁSSARO UM-UM TRONCO-POSP 3A-VT-RAIZ-PRS.CONT
'Each bird is on the trunk'

Six scenarios were produced for each experimental sentence, namely: (i) **standard distributive**, (ii) **partial distributive**, (iii) **extra-element**, (iv) **non-exhaustive**, (v) **collective,** and (vi) **uniqueness**. The **standard distributive** scenario (i) presented a prototypical distributive property. In other words, the elements of the sets of individuals and the subevents of the set of eventualities were distributed in one-to-one relationships, so that the number of relationships between the sets is greater than 2 and there is a differentiation between the sub-eventualities. The **partial distributive** scenario (ii) showed cardinality failure in the operation of relating individuals and eventualities, since the one-to-one relationship was violated in one of the subevents. The (iii) **extra-element** and (iv) **non-exhaustive** scenarios were complementary. In the first, there was an element of the group of individuals that did not participate in the avaiable subeventuality; in the second, there was a subeventuality that was not related to any individual in question. These scenarios tested the universality/exhaustion property of individuals and eventualities, respectively. The collective scenario (v) was composed of only one eventuality for which all individuals were applied, thus testing whether the property of differentiation between subeventualities was required by *sohoji-sohoji* in the course of processing. Finally, the **uniqueness** scenario (vi) included only one atom of the group of individuals that, in turn, participated in a single eventuality; therefore, it was tested whether a singular context was immediately discarded during the processing of the distributive numeral. For example, in Table (2) below, six scenarios are generated for sentence (15). Pictures were produced with reference to the picture database developed by Karajá designers within the framework of the PRODOCLIN|Karajá project.[3] Each picture was drawn with HB pencil on A4 sheet and scanned using 300-bpi scanner equipment. The

[3] Karajá Language and Culture Documentation Project, part of the Documentation Program for Indigenous Languages of the Indian Museum (FUNAI), supported by UNESCO.

scenarios were assembled from the scans with the help of Keynote software for Mac OS and exported in JPG format.

FIGURE 2 – Examples of experimental condition scenarios



Sentence: *Kruku sohoji-sohoji orà-tyreki ryireri*. 'Each bird is on the trunk'

The experiment was programmed in the E-prime software (Psychology Software Tools, Inc. – PST), a commonly used platform for presenting EEG-coupling language stimuli monitored by PyCorder software (Brain Vision LLC). Six versions of the experiment were developed in a Latin square format, so that no participant saw the same experimental sentences, but all the participants saw all the experimental conditions from different sentences.

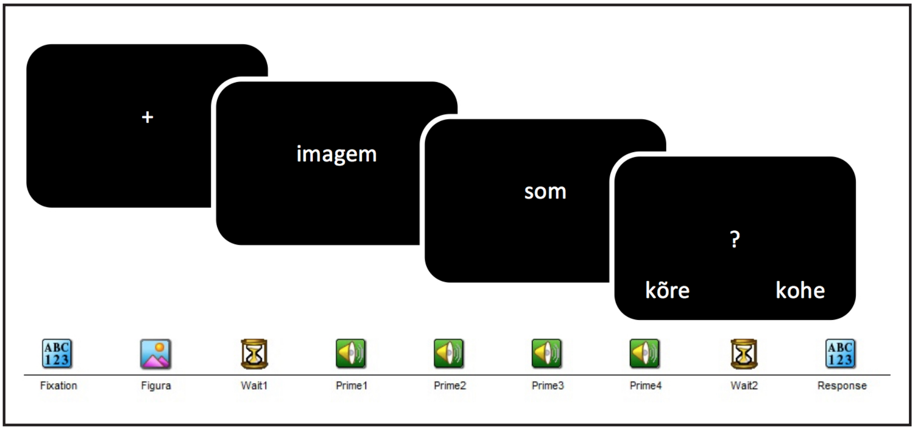
The lists were presented in both normal and inverse order to avoid list order effects in each version. Starting with the creation of experimental versions distributed among subjects – Latin square design – it is possible to ensure that the same subject never sees the same sentences, yet he/she is exposed to sentences of all conditions. In all, six lists/versions were programmed according to the distribution of the experimental materials in Latin square format. This experimental protocol is important to inhibiting extra-linguistic factors, such as memory, from interfering in sentence comprehension. The distribution of materials among the participants also helps the data to be more robust, since analysis between subjects dilutes subjective factors related to the individuals. In addition, the stimuli were pseudo-randomized within each list of each version, with the following restrictions: (i) experimental stimuli were separated by at least one distracting item and (ii) no experimental sentence was followed by another of the same condition.

The total set of materials consisted of 48 experimental sentences plus 48 distractors recorded in audio and 336 digitized and treated pictures. The audios were translated from Portuguese to Karajá with the help of two Karajá language teachers, conferred by a third speaker, and finally recorded by a Karajá consultant. Each experimental sentence was composed of four segments, complying with the Karajá language vocabulary standard, as described in more detail in example (16) below:

(16)	*Kruku*	*sohoji-sohoji*	*òra-tyreki*	*r-y-i-reri*
	bird	one-one	trunk-on	COP
	subject	distributive	locative	verb
	seg1	seg2	seg3	seg4

The participants were given the task of observing the picture and listening to the sentence that was presented by means of a monitor coupled simultaneously to a PC and an electroencephalography device. At the end of each <image, sound> pair, the task was to assess whether or not the materials were compatible with each other. The participants used a keyboard that had two keys highlighted in green and red. If the participant judged the sentence compatible with the picture, he pressed the green key. If he judged the sentence incompatible with the picture, he pressed the red key.

FIGURE 3 – Screen sequence presented to participants during the EEG experiment



As shown in Figure (3) above, the presentation of the stimulus set began with a cross in the center of the screen for 500ms, followed by the experimental picture/scene for 1500ms. Then the four audio segments were presented. The segments were controlled so that the average amount of time was 1500ms each. The total amount of time between segments was approximately 6000ms. After the sentence was heard, there was a 250ms pause, followed by the response screen, which had a 2000ms time-out, meaning that even if the participant did not press any of the answer buttons, a new data set would be introduced after two seconds. At the end of each experimental sequence, a blank screen lasting 1000 ms was presented before moving to a new <image, sentence> data set.

To determine whether or not there was compatibility between picture and sentence, the volunteers were instructed to push one of the two response buttons, which were balanced between the participants (right/left) to avoid side effects, i.e., to prevent those who already associated the right answer with the right side from benefiting from this association and vice versa. Response times were recorded by *E-prime* (*Psychology Software Tools, Inc*.) software.

EEG signals were continuously recorded by *PyCorder* (*Brain Vision LLC*) software from a system consisting of a system of 64 electrodes coupled in a plastic cap *(Brain Products GmbH*) in accordance with the 10-20 extended system (GOMES, 2014). The electrodes were allocated according to the international standard (*International System Locations*), including five electrodes along the median line (FPz, Fz, Cz, Pz, and Oz) and sixteen lateral/temporal channels, eight for each hemisphere of the brain (FP1/ FP2, F3/ F4, F7/ F8, C3/ C4, T3/ T4, T7/ T8, P3/ P4, and P7/ P8). In addition, another forty-three channels of the extended 10-20 system were used (AF3/ AF4, F1/ F2, F5/ F6, FC1/ FC2, FC3/ FC4, FC5/ FC6, FT7/ FT8, C1/ C2, C5/ C6, CP1/ CP2, CP3/ CP4, CP5/ CP6, TP7/ TP8, P1/ P2, P5/ P6, P7/ P8, PO3/ PO4, PO5/ PO6, PO7/ PO8, CB1/ CB2, and O1/ O2).

The electrodes were referenced online by the ground electrode, located in the frontal region, and later re-referenced by the mean of the mastoid channels (TP9/TP10 – left/right). Impedances were kept below 10 kΩ. The EEG was amplified and scanned at a frequency of 1000Hz. After recording, the data sampling frequency was maintained at 500Hz and the low-pass filters of 30Hz and high-pass of 0.1Hz were used, a method used by the ACESIN laboratory in accordance with international research standards (TANNER *et al*., 2016). Subsequently, ERP averages were generated for each subject per electrode at times ranging from 200ms to 1200ms in relation to the onset of the target stimulus. Times characterized by blinking or excessive muscle movement were automatically rejected by Vision Analyzer Software (Brain Products GmbH), a program used for filtering and averaging recorded data.

The extraction of ERPs is an electromagnetic cognitive evaluation tool whose application has been extended to linguistic research in the neuroscience of language field since the 1980s, having been developed from the EEG invented by German psychiatrist Hans Berger and used with humans since 1929.[4] The ERP technique is regarded as very accurate because it registers the central nervous system’s responses to motor or sensorial stimulation with millisecond precision and is thus considered an online measurement technique. The EEG provides a continuous temporal sampling of the brain’s electrical activity. Such sampling may be

[4] Like the methods used in experimental psycholinguistics, the ERP technique is also non-invasive, thus allowing it to be used in studies on cognition and language in different human groups, of different age groups, with or without pathologies, and without harm or strain for the participant. To execute the technique, all that is needed is the equipment and software for the gauging, treatment, and statistics of the components that one wishes to investigate.

related to different types of linguistic phenomena and is monitored with the help of specific software by an EEG device attached to a computer that presents the experiment's participants to a certain pre-established linguistic task. After picking up the electrical signals produced in the cortex and monitoring them through the EEG, they must be treated (filters are applied) so that the data can be analyzed (GESUALDI; FRANÇA, 2011). Filtering is an essential, inseparable part of ERP research.

> ERPs are composed of a sequence of waves characterized by their latency, amplitude, and polarity. ERP usually has an instantaneous value of 10 to 1000 times less than the background EEG and, therefore, cannot be visualized. To be visualized, the average of several times needs to be performed[5]. This procedure is justified by assuming spontaneous EEG as a zero mean Gaussian white noise and ERPs as the only responses that are actually synchronized with the stimulus. In this way, the effect of averaging is to increase the signal-to-noise ratio (SNR), thus allowing the visualization of the specific effect of the stimulus, in this case, linguistic. (MARQUES, 2011, p. 30)

Thus, averaging is a mathematical treatment of the data used to consolidate the average of the same segment around a given time-coupled moment for all of the stimuli. This treatment aims to exclude random noise that is concurrent with the signal related to the event being studied, such as those resulting from muscular movements, eye blinks, or electrical interference caused by some equipment in the testing room. Since the noise is random, the median after averaging would tend to be zero. Therefore, at the end of this "cleaning" operation, only the signal related to the linguistic investigation will emerge, meaning a large average (or comprehensive average) of the event windows linked to the display time of the stimuli developed for this investigation (LAGE, 2005).

The N400 component was discovered by Kutas and Hillyard in the late 1970s. The researchers developed an experiment to elicit the P3 component (already widely investigated at the time for word recognition) by investigating the role of context in word recognition in sentences such as "eyebrows" in "He shaved off his mustache and

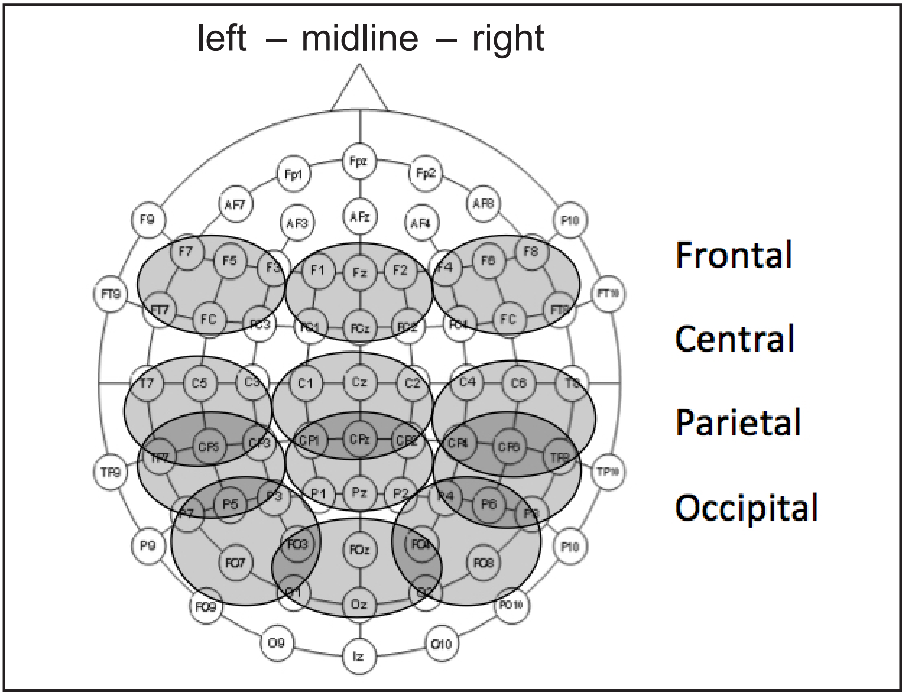
[5] Times or events are working windows in the temporal continuum of the EEG demarcated for the study.

eyebrows." The results showed a small negative component rather than a positive one as initially expected. To more clearly observe this effect, the authors applied a new experiment with sentences containing a word that, instead of being unanticipated by the context, was semantically anomalous, sentences like "He shaved off his mustache and city" and "I take my coffee with cream and dog". In the second experiment, the negativity effect was more pronounced, since it elicited very high negative polarity peaks around 400 ms in the course word presentation (KUTAS; HILLYARD, 1980). Since then, this has been the most studied component in the language field, as it is mainly related to the semantic content of the words per se and lexical access theories, since it seems to provide clues on how words are stored (FRANÇA *et al*., 2008). In Brazil, the study of the N400 component was pioneered in the doctoral thesis of Aniela Improta França (FRANÇA, 2002) entitled *Linguistic concatenations: a study of different cognitive modules in acquisition and in the cortex*. Topographically, N400 tends to exhibit increased activation in the left central and parietal regions of the brain.

Another component studied by language neuroscience, P600, was more recently elicited by Osterhout and Holcomb (1992) and, since then, has been related to syntactic anomalies and incongruities, and can be elicited in sentences of type **The broker persuaded to sell the stock*, in contrast with well-formed sentences like *The broker hoped to sell the stock.* Gouvea *et al*. (1995) also found evidence of high amplitude in the P600 component for sentences that generate temporally incorrect analyses (known in psycholinguistics as garden path sentences). P600 has also been elicited in non-linguistic anomalies in music (PATEL *et al*., 1998) and mathematics (MARTÍN-LOECHES *et al*., 2006). In the languages area, in addition to sentence violations, the P600 component also seems to reflect syntax violations, such as violation of number, time, gender, and case agreement, as mentioned by Gouvea *et al*. (1995) and investigated by Gomes (2014), Gomes and França (2015), and Gomes and França (2013). The Late Positive Component (or Late Positive Complex) has already been elicited in studies on the quantification of statements (JIANG *et al*., 2009; KAAN *et al*., 2007), which appear to show that quantified sentences present more complex processing, because the effects of quantificational incongruence are captured at a later time than other language phenomena. Kim and Osterhout (2005) evaluate that this component can be observed in items whose semantic incoherence could lead to a syntactic reanalysis.

With regard to the analyses presented below, the N400 and P600 components of interest were visually inspected and their peaks were automatically recorded by Vision Analyzer. The peak amplitude and voltage of the components of interest for each of the channels were captured in all segments and conditions. The N400 component was inspected within a 350-550ms window (capturing wide negativity), and the P600 component was inspected within a 550-800ms window (capturing wide positivity) after the onset of the target stimulus. Grand averages were formed by the ERP data of each participant's interest.

FIGURE 4 – Regions of interest based on anatomical proximity



The dependent measures, i.e., the amplitude and latency of the N400 and P600 components, were analyzed by analysis of variance (ANOVA) and parametric tests (T-test). ANOVAs were applied between the *Condition* (distributive, partial distributive, collective, extra-element, non-exhaustive, and uniqueness), *Laterality* (left hemisphere, right hemisphere, and midline), and *Anteriority* (frontal, central, parietal, and occipital regions) factors. Parametric tests were performed to check any significant difference between the standard distributive condition and other experimental conditions. Therefore, we performed four distinct analyses based on the ERP data:

i. ANOVA of the peak amplitudes of the waves for the N400 component by condition, laterality, and anteriority. The t test contrasting the means obtained with those of the standard distributive condition.
ii. ANOVA of the peak amplitudes of the waves for the P600 component by condition, laterality, and anteriority. The t test contrasting the means obtained with those of the standard distributive condition.

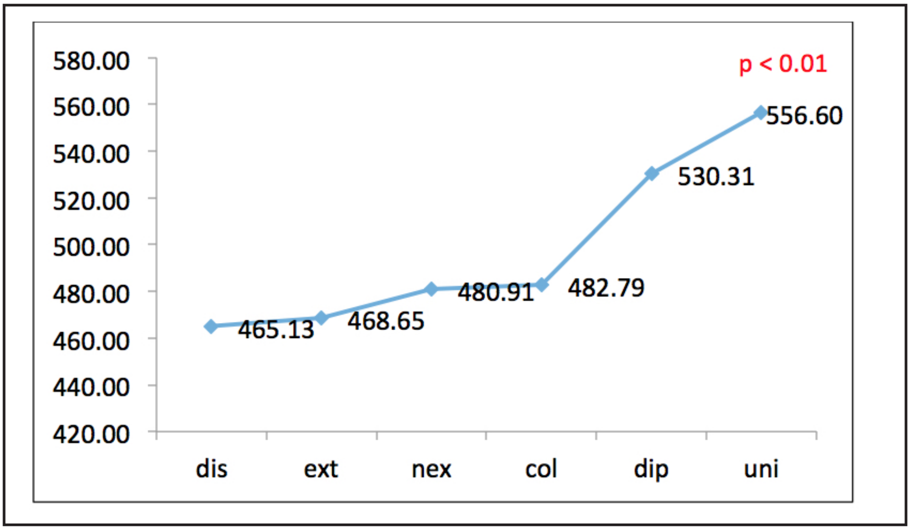
Statistics based on the N400 and P600 components are the most discussed measures in the neuroscience of language literature, and therefore represent the best starting point for an experiment of this nature. Moreover, studies that investigated different aspects of the quantification of statements elicited N400 in quantified sentences, evidencing a higher initial processing cost for statements that involve procedures such as counting and individuation (CHIARELLI *et al*., 2011). Late positivity effects were also observed, interpreted by Jiang, Tan & Zhou (2009) as an effort to integrate quantifying elements into sentences that present quantificational inadequacy. Therefore, for the analyses presented below, we extracted the amplitude of the N400 and P600 peaks in all experimental conditions from the large means of segments 2 and 3 of the manipulated sentences.

4.3 Behavioral Data

An analysis performed by means of a t test of participants' reaction times (RTs) allowed us to observe which conditions presented the highest average decision time in relation to the standard distributive condition. The results show that only the **uniqueness** condition ($t = -2.585$, $df = 21.182$, p-value $= 0.0172$) elicited significantly higher

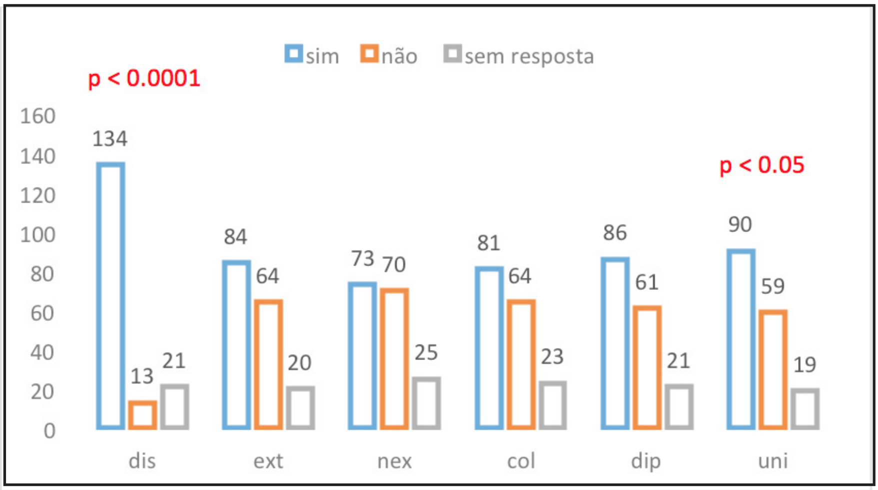
RTs than the distributive condition, which suggests a higher cognitive cost for the decision in this condition. The analysis of raw RTs shows that the uniqueness scenario caused the greatest feeling of strangeness during the pairing task with a sentence containing the distributive numeral. Other conditions did not present significant differences, as can be seen in Graph (1) below.

GRAPH 1 – Raw RT per experimental condition



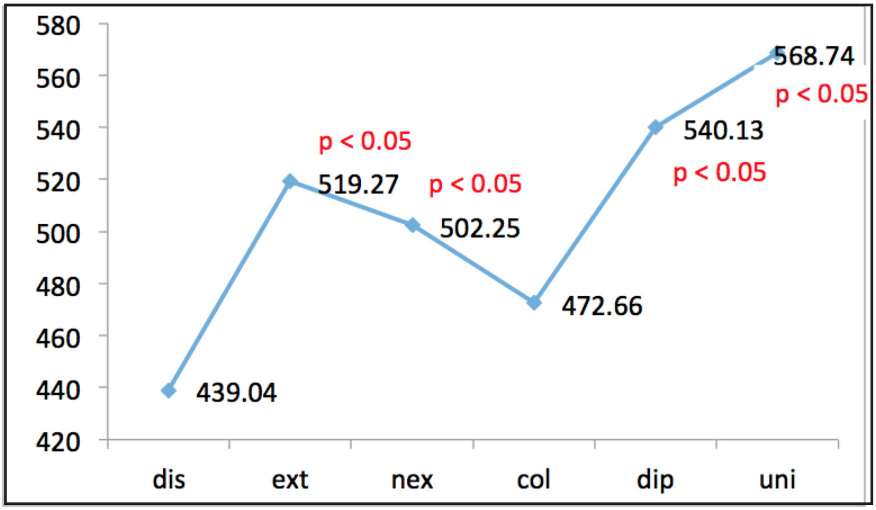
In general, there is a tendency to accept all conditions as compatible scenarios for *sohoji-sohoji* sentences, as shown in Figure 2 below. However, after performing chi-square tests for each condition, it was observed that only the distributive (X-squared = 99.5986, df = NA, p-value = 0.0004998) and uniqueness (X-squared = 6.4497, df = NA, p-value = 0.02249) conditions showed a significant difference between "yes" and "no" answers. The fact that the distributive condition has the highest acceptance rate confirms our expectations since this is the best scenario for the investigated experimental sentences. The acceptance of the uniqueness condition seems to be the result of a cardinal reading attempt of the reduplicated numeral, because, as we will see in the bioelectrical data, this condition was the one that had the highest cognitive cost during processing. Although the other conditions did not capture a significant effect in this test (p> 0.05), we observed that more reflex measures, such as RTs and bioelectrical potentials, reveal there are differences in processing between the different conditions and the standard distributive condition. The 2000ms time-out generated some "no response" items that were excluded from the analyses.

GRAPH 2 – Response rates by condition



Since the rates for "yes" responses were higher in all conditions, the t-test was again applied between conditions to examine only the RTs for the scenarios accepted by the participants, excluding scenarios deemed to be incompatible. **Partial distributive** and **uniqueness** conditions continued to be those with the highest averages, in which case both presented significant differences. However, **non-exhaustiv**e and **extra element** conditions also attained significantly higher times. The **collective** condition did not present significant RT differences when accepted by the speaker; in this case, the mean was statistically similar to the standard distributive condition. The p-values and averages can be seen in Graph 3 below.

GRAPH 3 – RT of the accepted pairings per experimental condition



These results highlighted the importance of the chronometric data for our analysis, since the acceptability indexes did not prove to be a decisive factor among all the scenarios. Furthermore, although the deviant uniqueness condition has been accepted as a valid scenario for the experimental sentence, the RT analysis shows that this acceptability was performed in a significantly longer period of time than for the standard distributive condition. Without decision time analysis, it would also be impossible to capture the significant distinction between the other experimental conditions, since acceptability remained constant for all scenarios. However, we still cannot get evidence from the RTs that there is a difference in processing between non-exhaustive and extra-element conditions, nor capture if there is a greater cognitive effort to accept the condition uniqueness, as expected.

The behavioral results presented in this section were captured after sentence processing, i.e., offline. We also captured the electrophysiological responses through the ERP extraction technique for online data analysis. The purpose of this technique was to observe the processing of experimental sentences from moment to moment during the temporal course of sentence comprehension. The results of this method will be presented in the next subsection, point out new reflections regarding the understanding of the processing of the *sohoji-sohoji* distributive, and reveal interesting data on how the acceptability of different scenarios is given.

4.4 ERP Data

The three-way ANOVA (within subjects) model was first used for statistical analysis, which included the experimental condition, laterality, and ROI factors, with correction of the values presented (Greenhouse-Geisser). After the data had been treated, the N400 component showed a significant main effect for the condition factor **(F (5.25) = 7.56 p <0.0000003)** in the second segment of the experimental sentences, corresponding to the *sohoji-sohoji* reduplicated numeral. Conversely, the P600 component elicited a main effect for the condition factor **(F (5.25) = 14.7 p <0.0017565)** only in segment 3, corresponding to the locative postpositional (PP) phrase. No major effects were found for the ROI (F (11.55) = 1.19 p <0.3472345) or laterality (F (2.10) = 1.90 p <0.2238272) factors in the P600 component nor for the ROI (F (11. 55) = 1.41 p <0.195080 p <0.2820406) or laterality (F (2.10) = 2.55 p <0.1680510) factors in the N400 component. Nor did we find any interaction among the mentioned factors. Therefore, it seems that only the condition factor appears to influence the analyses between the results. Consequently, an analysis was carried out including all ROIs for this article. In the future, other analyses may be performed to confirm the trends presented here.

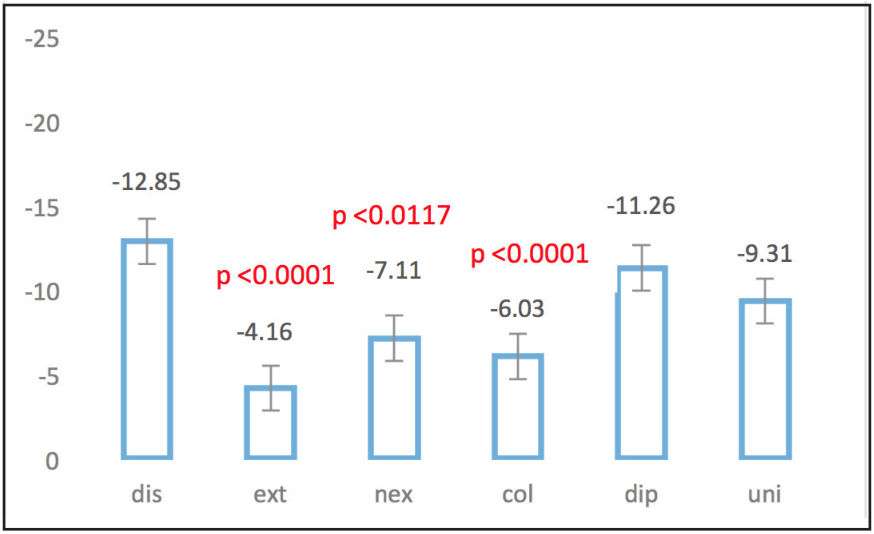
Subsequently, a t test was applied that contrasted the standard distributive condition with the other experimental conditions. Presented below are the results for the amplitude measurements (μV) of the ERP data of the Grand Averages for the N400 and P600 components.

4.4.1 N400 component – negativity after onset of *sohoji-sohoji*

The N400 component, ascertained after the start of the segment containing the *sohoji-sohiji* reduplicated numeral, presented significant results when comparing the standard distributive condition (-13 μV) with the **non-exhaustive** (-7 μV, t (120) = 2.56 p <0.0117), **collective** (-6 μV, t (120) = 4.88 p <0.0001), and **extra element** (-4 μV, t (120) = 6.13

p <0.0001) conditions. The uniqueness and partial distributive conditions did not present significant differences.

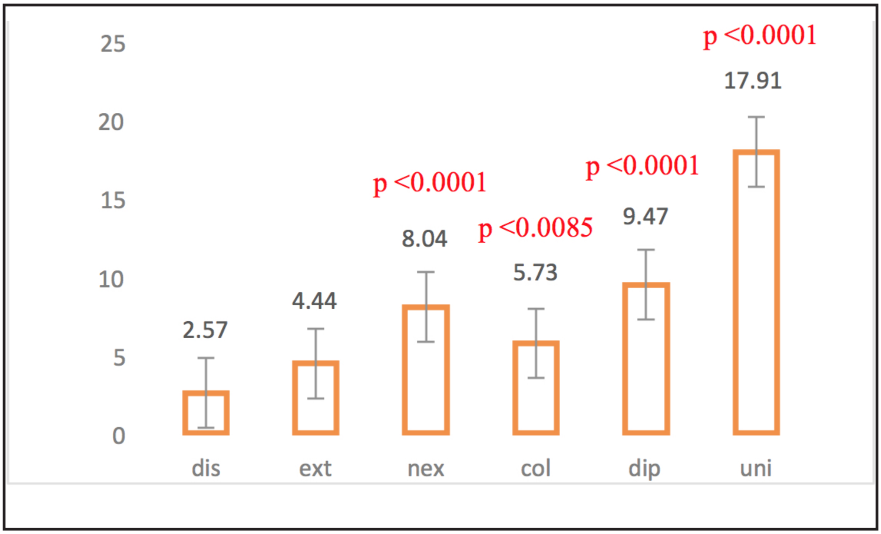
GRAPH 4 – N400 component amplitude averages per condition



4.4.2 P600 component – late positivity after onset of locative PP

The P600 component, ascertained after the start of the segment containing the locative PP of the eventuality, elicited a significant difference between the **uniqueness** (18 µV, $t(118) = 17.75$ $p < 0.0001$), **partial distributive** (9 µV, $t(118)=4.91$ $p< 0.0001$), **non-exhaustive** (8 µV, $t(118)=4.17$ $p < 0.0001$), and **collective** (6 µV, $t(118)=2.67$ $p< 0.0085$) conditions when compared to the standard distributive condition (3 µV).

GRAPH 5 – P600 component amplitude averages per condition



4.5 Discussion of results and analyses

The aim of this experiment was to search for a microscopic understanding of the processing of the *sohoji*-*sohoji* distributive, using previously developed contexts that were provided by scenarios controlled by a specific design, as described at the beginning of this section. The method made it possible to analyze electrophysiological data captured by the EEG during sentence processing, measured online, that would reflect the cognitive cost during sentence processing when paired with different quantitative scenarios. The data were then filtered for extraction and analysis of the amplitudes (µV) captured from the delineation of N400 (250-500ms window) and P600 (550-800ms window). The results indicate that the acceptability of different experimental conditions in behavioral data does not reflect the cognitive cost for pairing between the different scenarios and the reduplicated numeral, since the participants tended to accept all the conditions. However, there was a tendency for slower reaction times (RTs) for scenarios deviating from the distributive pattern, namely, those that allowed intermediate and non-pluralized distributive readings. Therefore, the results confirm the earlier hypothesis that the processing of cumulative readings, although possible, would be more costly than non-cumulative sentences when there is an element that delineates the quantificational scope, in this case the *sohoji-sohoji* reduplicated numeral. As discussed below, online data suggest that the properties of cumulative property, distributive property, and cardinality exert effects during sentence processing with *sohoji-sohoji* in Karajá. This evidence is in line with current theories on quantification processing and the semantics of distributional numerals, which have been observed in other natural languages.

With regard to ERP amplitude data, N400 elicited lower amplitudes for the **collective** and **extra-element** conditions in segment 2, corresponding to the *sohoji-sohoji* reduplicated numeral. The P600 component elicited higher amplitudes for the **non-exhaustive**, **collective**, **partial distributive**, and **uniqueness** conditions in segment 3, corresponding to the locative PP. These results seem to reflect the cognitive cost of processing the reduplicated numeral against the different scenarios presented during the experiment. The N400 component reveals that the prototypical distributive scenario, that for which there is an individual for each eventuality, has the highest average amplitude, followed by the partial distributive condition. Such evidence indicates that the processing of *sohoji-sohoji* would engage the individuals' cognitive system in a distributive operation requiring a higher processing cost at the reflex level. As for the P600 measurement, registered in the locative segment following *sohoji-sohoji*, the conditions with larger amplitudes are those that break the event-distributive reading, revealing that it is after the start of the pairing procedure of the element in the eventuality that measurements emerge that would show which properties would be more easily processed in distributive numeral sentences.

The **uniqueness** condition was the one that elicited higher mean values in both the P600 (18 μV) component and the raw behavioral data (557 ms) and filtered by acceptability (568 ms), which suggests the reduplicated numeral requires the pluralization of the statement with respect to the relations established by the individuals and available eventualities, in agreement with what is expected by distributive elements.

The **partial distributive** condition also presented a high amplitude average for the P600 component (9 µV) and was also the second highest measurement of the reaction times filtered by acceptability (540 ms). These results indicate that intermediate readings, although accepted, present a high processing cost. In this case, as *sohoji-sohoji* is formed from the reduplication of the number one, these results can also be analyzed as coming from the numerical cardinality failure presented by the presented scenario, i.e., a problem in the counting operation.

The **non-exhaustive** condition elicited the P600 component (8 µV) and significant reaction times filtered by acceptability (502 ms). In this condition, an element of eventuality was not in an R relation with any individual in the scenario. The high processing cost shows that the interpretation of sentences with *sohoji-sohoji* requires that the scenario subevents preferentially operate under the scope of a universal quantifier.

Despite eliciting the P600 component (6 µV), the **collective** condition presented lower mean values than did the other conditions mentioned above. Moreover, no effect was captured in the behavioral data. This effect could possibly have been potentiated if events were verified with transitive verbs that allowed a collective and not only a locative reading of the event. This hypothesis is supported by the fact that the participants accepted the collective scenario presented after full sentence processing without significant cognitive costs, since no significant effects on the RTs were found.

The **extra element** condition did not show significant differences in the P600 component, i.e., a scenario in which an individual who did not participate in the event in distribution did not affect the positive amplitude (4 µV). However, in the behavioral results filtered by accuracy (519 ms), this condition presented significant results, indicating that, on a more reflexive, online level, the universality over events is less costly than the universality over individuals during the *sohoji-sohoji* distributive operation. The difference found in the RT may have been the effect of a reinterpretation of post-sentence processing.

Regarding the significant results found in N400 component for the **extra element** (-4 μV), **non-exhaustive** (-7 μV), and **collective** (-6 μV) conditions, one proposal is that, under these conditions, there is a greater delay for engagement of the memory system in computing the distributive property, since lower means reflect less cognitive effort. Therefore, it is appropriate to propose that these results reflect that the semantic incompatibilities of universality and differentiation present in these scenarios would lead the cognitive system to initially not engage in the verification of compatibility of such scenarios, procrastinating verification to a later processing moment, as verified by the results in the P600 component and in the behavioral data. The negativity elicited by the N400 component for the **standard distributive** (-13 μV) and **partial distributive** (-11 μV) conditions seems to be a result of the incrementality effect of the quantifiers, which would be interpreted in stages, as is the case with negation. (URBACH; KUTAS, 2010; e URBACH; DELONG; KUTAS, 2015). The negativity elicited for the **uniqueness** condition (-9 μV) may be a consequence of semantic implausibility in comparing a non-pluralized scenario with a distributive sentence, which would generate a participant's neurophysiological response reflecting the shock in processing a sentence operated by a distributive numeral after being exposed to a singular scenario, which should be immediately discarded.

5. Conclusion

This paper aimed to demonstrate how the availability of ever more sensitive techniques for discovering how the study of language processing can enrich traditional field research and formal language analysis. In this interdisciplinary perspective, the EEG experiment to extract ERPs aimed to investigate how participants would process distributive sentences with *sohoji-sohoji* reduplicated numerals when exposed to scenarios that presented intermediate readings, rather than prototypically distributive ones.

The results demonstrate that the N400 and P600 components were elicited when there were incompatibilities between scenario and sentence. The reanalysis of conditions that violated plurality, cardinality, differentiation, or exhaustiveness was captured by a wide positivity in the P600 component, elicited from elements that occurred as locative phrases of events. At a previous time, a wide negativity effect on the N400 component was captured under conditions that engaged participants in a quantitative procedure (standard distributive and partial distributive) or immediately dismissed the plausibility between the scenario and the distributive sentence (uniqueness).

These results complement those found in offline chronometric measurements that point out that *sohoji-sohoji* could or would act as an ambiguous quantifier, allowing more than one reading to be licensed for the sentence (event-distributive or participant-distributive readings) or license different scenarios of verification for the same reading, in a perspective close to that proposed by Cable (2014), who proposes a semantic denotation capable of accounting for the different verification scenarios inherent to the distributive quantifiers in Tlingit whose resolution will depend mainly on the context.

In addition, if Choe's (1987) proposal that the distributive operator is a universal quantifier, which has both a constraint set and a scope set, is taken into account, the *sohoji-sohoji* reduplicated numeral as a default would put the eventuality in the set in restriction, i.e., the event-distributive reading would be less costly. This is because the universal operator (∀) should be applied to the members of the set of eventualities. In other words, the non-exhaustiveness of the elements of the sentence's eventuality would be the reading that would present higher cognitive cost, because it needs reanalysis, as confirmed in the neurophysiological data.

Distributive property is a knowledge that is part of human cognition and found in different natural languages through various morphological and syntactic mechanisms. It is hoped that this work will give visibility to the phenomenon of numeral distributions found in Amerindian languages that may contain relevant properties for the study of areas pertinent to linguistics and, therefore, knowledge of the faculty of human language.

Bibliography

CABLE, S. Distributive numerals and distance distributivity in tlingit (and beyond). *Language*, Washington, Linguistic Society of America, v. 90, n. 4, p. 562-606, 2014.

CHIARELLI, V.; YAGOUBI, R.; MONDINI, S.; BISIACCHI, P.; SEMENZA, C. The syntactic and semantic processing of mass and count nouns: an ERP study. *PLoS ONE*, San Francisco, CA, v. 6, n. 10, p. 1-15, 2011.

DOWTY, D. Collective Predicates, Distributive Predicates, and All. In: ESCOL, 3rd, 1987, Columbus. *Proceedings*... Columbus: Ohio State University, 1987.

FRANÇA, A. I. *Concatenações linguísticas*: estudo de diferentes módulos cognitivos na aquisição e no córtex. 2002. Tese (Doutorado em Linguística) – UFRJ, Rio de Janeiro, 2002.

FRANÇA, A. I.; LEMLE, M.; GESUALDI, A.; CAGY, M.; INFANTOSI, A. F. C. A neurofisiologia do acesso lexical: palavras em português. *Revista Veredas*, v. 2, p. 34-49, UFJF, Juiz de Fora, 2008.

GESUALDI, A. da R.; FRANÇA, A. I. Event-related brain potentials (ERP): an overview. *Revista LinguiStica*, UFRJ, Rio de Janeiro, v. 7, n. 2, p. 24-42, 2011.

GIL, D. *Distributive numerals*. 1982. Dissertation (Ph.D.) – University of California, Los Angeles, 1982.

GOMES, J. N. *Investigating the dissociation between n400 and p600 effects on the syntax semantics interface*: an ERP study. 2014. Tese (Doutorado) – Universidade Federal do Rio de Janeiro, Faculdade de Letras, Rio de Janeiro, 2014.

GOMES, J.; FRANÇA, A. I. A técnica de ERP: investigando a assimetria sujeito-objeto na interface sintaxe-semântica com EEG. *Letras de Hoje* (Online), Pontifícia Universidade Católica do Rio Grande do Sul, v. 4, p. 152-163, 2015.

GOMES, J. N.; FRANÇA, A. I. The ample semantic scope of minute language computations: an ERP study of words in Portuguese. *Revista da ANPOLL*, Anpoll, v. 34, p. 309-346, 2013.

GOUVEA, A. C.; PHILLIPS, C.; KAZANINA, N.; POEPPEL, D. The linguistic processes underlying the P600. *Language and Cognitive Processes*, Taylor & Francis Online, v. 25, p. 149-188, 1995.

HEIM, I. *The semantics of definite and indefinite noun phrases*. 1982. Thesis (PhD) – GLSA, University of Massachusetts, Amherst, 1982.

JIANG, X.; TAN, T.; ZHOU, X. Processing the universal quantifier during sentence comprehension: ERP evidence. *Neuropsychologia*, Elsevier, v. 47, p. 1799-1851, 2009.

KAAN, E.; DALLAS, A. C.; BARKLEY, C. M. Processing bare quantifiers in discourse. *Brain Research*, Elsevier, v. 1146, p. 199-209, 2007.

KIM, A.; OSTERHOUT, L. The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, Elsevier, v. 52, p. 205-225 2005.

KRATZER, A. On the plurality of verbs. In: DÖLLING, J.; HEYDE-ZYBATOW, T. (Ed.). *Event structures in linguistic form and interpretation*. Berlin: Mouton de Gruyter, 2007.

KRATZER, A. The event argument and the semantics of verbs. 2003. Manuscrito. Disponível em: http://semanticsarchive.net.

KRIFKA, M. The origins of telicity. In: ROTHSTEIN, S. (Ed.). *Events and grammar*. Dordrecht: Kluwer, 1998. p. 197-235.

KUTAS, M.; HILLYARD, S. Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*, American Association for the Advancement of Science, v. 207, p. 203-205, Jan. 1980.

LAGE, A. C. *Aspectos neurofisiológicos de concatenação e idiomaticidade em português do Brasil*: um estudo de potenciais bioelétricos relacionados a eventos linguísticos (ERPs). 2005. Tese (Doutorado em Linguística) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.

LYONS, J. *Semantics*. London: Cambridge University Press, 1977. 2 v.

MARQUES, F. *Eletrofisiologia da idiomaticidade em indivíduos com Síndrome de Asperger*: um estudo de ERPs. 2011. Dissertação (Mestrado) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2011.

MARTÍN-LOECHES, M.; CASADO, P.; GONZALO, R., DE HERAS, L.; FERNÁNDEZ-FRÍAS, C. Brain potentials to mathematical syntax problems. *Psychophysiology*, Wiley Online Library, v. 43, p. 579-591, 2006.

MAY, R. *The grammar of quantification*. 1977. Dissertation (Ph.D.) – MIT, 1977.

MAY, R. *Logical form*: its structure and derivation. Cambridge, Massachusetts: MIT Press, 1985.

MULLER, A. Distributividade: o caso dos numerais reduplicados em Karitiana. *Cadernos de Estudos Linguísticos*, Campinas, v. 54, n. 2, p. 223-243, 2012.

OLDFIELD, R. C. The assessment and analysis of handedness: the Edinbugh inventory. *Neuropsychologia*, Elsevier, v. 9, p. 97-113, 1971.

OSTERHOUT, L.; HOLCOMB, P. J. Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, Elsevier, v. 31, n. 6, p. 785-806, 1992.

PATEL, A. D.; GIBSON, E.; RATNER, J.; BESSON, M.; HOLCOMB, P. J. Processing syntactic relations in language and music: an event-related potential study. *Journal of Cognitive Neuroscience*, MIT Press, v. 10, n. 6, p. 717-733, 1998.

SILVA, C. O. da. *O processamento do numeral distributivo sohoji-sohoji na língua karajá*. 2016. Tese (Doutorado em Linguística) – Faculdade de Letras da Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2016.

TANNER, D.; NORTON, J. J.; MORGAN-SHORT, K.; LUCK, S. J. On high-pass filter artifacts(they're real) and baseline correction (it's a good idea) in ERP/ERMF analysis. *Journal of Neuroscience Methods*, Elsevier, v. 15, n. 266, p. 166-170, Jun. 2016.

TUNSTALL, S. *The interpretation of quantifiers: semantic and* *processing*. 1998. Dissertation (PhD) – University of Massachusetts, Amherst, 1998.

URBACH, T. P.; DELONG, K.; KUTAS, M. Quantifiers are incrementally interpreted in contexto, more than less. *Journal of Memory and Language*, Elsevier, v. 83, p. 79-96, 2015.

URBACH, T. P.; KUTAS, M. Quantifiers more or less quantify on-line: ERP evidence for partial incremental interpretation. *Journal of Memory and Language*, Elsevier, v. 63, p. 158-179, 2010.