NOTES ON SPEAKERS' INTUITIONS AND JUDGMENTS IN LINGUISTIC THEORY AND PSYCHOLINGUISTICS: THE CASE OF Q-EXPRESSIONS IN BRP[1]

Mercedes Marcilese[2]
Erica dos Santos Rodrigues[3]

mmarcilese@gmail.com
ericasr@puc-rio.br

ABSTRACT: The data on which linguists base their theories typically consist of intuitive judgments of the well-formedness of utterances in a given language. Grammaticality and acceptability judgments are one of the most widespread data-collection methods used to test theoretical linguistic claims. However, the validity criteria underpinning this kind of method have been the target of an interesting debate in the literature. This paper aims at addressing this topic by means of a comparison between linguistic proposals—mainly based on informal judgments—and experimental results concerning the interpretation of the Brazilian Portuguese universal quantifiers *cada*, *todo,* and *todosos.* The set of experimental results suggest that several methodological aspects—mainly the presence or absence of time pressure during the performance of experimental tasks, as well as the type of task itself—may affect the processing, the interpretation and, as a consequence, the judgment of sentences containing quantified expressions. We assume that both linguists and psycholinguists are exploring the same cognitive system, albeit with different tools.
Keywords: Acceptability/Grammaticality Judgments; Linguistic Theory; Psycholinguistics; Universal Quantifiers; Brazilian Portuguese.

Grammaticality or acceptability judgments are a methodology traditionally used by generative linguists in the research on syntactic and semantic phenomena. The validity criteria underpinning this kind of method have been the subject of a deep debate in the area (Schütze 1996; Gibson, Piantadosi, and Fedorenko 2012; Gibson and Fedorenko 2010a, 2010b; Bresnan 2007; Sprouseand Almeida 2010, 2012a; Schützeand Sprouse 2013; among several others). The main criticisms concern the validity and reliability of grammaticality judgment tests as a means of making inferences about linguistic competence. Several methodological aspects have been criticized, such as the reduced number of participants, the small number of structures of interest, and the lack of control regarding the order of stimulus presentation (Gibson, Piantadosi, and Fedorenko 2012; Gibson and Fedorenko 2010a, 2010b). There are also objections to judgment data since judgment tasks usually require metalinguistic awareness of language along with decision-making by the speakers, which would undermine their external validity (Bresnan 2007). In response to these criticisms, some authors advocate in favor of the use of informal acceptability judgment experiments as the primary source of linguistic data (Sprouse and Almeida 2010, 2012a). In this paper,we address this topic by means of a comparison between linguistic analyses based on informal acceptability judgments and experimental results concerning the interpretation of the Brazilian Portuguese (BrP) universal quantifiers (Qs)*cada*, *todo,* and *todos os* by native speakers (Marcilese and Rodrigues 2014a, 2014b; Rodrigues and Marcilese2014b, submitted). In the next sections, we briefly discuss the role of the speaker's intuitions in Linguistic Theory, as well asthe psychological side of acceptability/grammaticality judgments. We also contrast the description of the Q-expressions provided by linguistic theory—primarily based on informal judgments[4]—and the characterization that arises from the experimental data analysis. Our final remarks are aimed at discussing the relevance of the use of both formal and informal judgments in order to investigate linguistic phenomena.

---

[4] The expression "informal judgments" is used here in the sense proposed by Sprouse and Almeida (2011, 2012b). The meaning of the adjective "informal" is defined when traditional judgment gathering is compared with the standards of experimental cognitive science.

1. SPEAKERS' INTUITIONS AND JUDGMENTS IN LINGUISTIC THEORY

According to Levelt (1972:20): "the essence of the relation between linguistics and psychology is the "interfacing' of competence and performance." Linguistic competence is defined as the "creative language capacity of the language user",that is, a—mainly unconscious—knowledge that enables the speaker to produce and understand a potentially unlimited number of sentences. Performance, in turn, is characterized as the actual use of this knowledge in several kinds of language activities, such as speaking, hearing, reading, and writing.The distinction between these terms that refer to the speaker's knowledge of language— the competence—and to the speaker's use of this knowledge—the performance—were proposed by Chomsky in *Aspects of the Theory of Syntax* (1965), in his attempt to stress that a generative grammar aims to mirror the speaker's competence, by capturing the creative nature of human languages.

Grammaticality judgments are one of the most widespread data-collection methods used to test theoretical linguistic claims or, in other words, used in order to sketch linguistic competence. In judgment tasks, speakers are presented with a set of linguistic stimuli and they must determine whether—and/or the extent to which—the stimulus could be legitimately generated by the grammar of a given language.The data on which linguists base their theories typically consist of intuitive judgments of the well-formedness of utterances in a certain language.

On the other hand, the study of linguistic performance must be considered "the interaction of a variety of factors, of which the underlying competence of the speaker-hearer is only one" (Chomsky1965: 4). These"other" performance factors include psychological variables such as attention, memory span, world knowledge,etc.

As statedby Tremblay (2005), the use of judgments tasks in linguistic theory is crucial in order to (i) assess the speakers' evaluation of sentences that rarely occur in spontaneous speech; (b) obtain negative evidence on strings of words that are not part of the language; (c) distinguish performance problems (e.g., slips, attraction errors, unfinished utterances, etc.) from production of grammatical structures; and (d) isolate the structural properties of the language. Both elicited production tasks and naturalistic data collection alone may not be sufficient to achieve these goals.

The distinction between a grammaticality judgment andan acceptability judgment is more than just a terminological issue.Moreover, the term *grammaticality* itself has been considered controversial; speakers' judgments concerning whether a particular string of words

is or is not a possible utterance of their language are, in fact,*percepts*[5](Schütze and Sprouse 2011). Grammar, defined as a mental construct, is not accessible to conscious awareness. In this sense, the term "acceptability"would be more adequate than "grammaticality." An acceptability judgment is an indirect measurement method used to evaluate linguistic hypothesis about grammatical constructs, including formal representations and operations.

In contrast to psycholinguists, linguists are usually not concerned with online processes, with how representations are constructed or retrieved in real-time tasks. This does not entail, however, that acceptability or grammaticality judgments are not psychologically real.This misconception is directly related to the competence/performance distinction, which is usually mentioned for justifying a division of labor between linguists and psycholinguists, and has contributed to the idea that the two groups investigate necessarily different objects. In Phillips and Wagers (2007:740) words: "the data that linguists and psycholinguists collect and the theories that they develop based on those data are all 'psychological', in the sense that they aim to explain some aspect of human cognitive abilities." It is an empirical hypothesis, however, whether the cognitive system that linguists investigate is distinct from what the psycholinguists are concerned with.

Schütze (1996) claims that linguists are generally not concerned with methodological issues and offers a detailed discussion of objections to the use of grammaticality judgments. Despite the fact that the author himself argues that the collection of grammaticality judgments is legitimate, he points out problems with certain methods of eliciting and interpreting those judgments. In particular, Shütze points out thelack of standard experimental controls in the data collection: number and order of trials, absence of fillers, number and naiveté of subjects, task presentation, statistical analyses, etc.

Gibson and Fedorenko (2010 a, b) share Shütze's criticism, and they also highlight the weak quantitative standards in linguistics research. The authors advocate that multiple items and multiple naïve participants should be evaluated and that quantitative analysis methods are necessary. In response, Sprouse and Almeida(2011, 2012b; among others) estimate that the maximum replication failure rate for the informal results in syntax research is 2% (i.e. the empirical foundation of the theory is at least 98% replicable). These results suggest that the extensive use of informally collected judgments in syntax has not led to theories constructed upon faulty data. According to Sprouse & Almeida, there are several reasons to adopt experimental methods, but the inadequacy of the empirical foundation of the theories is not one

---

[5] A *percept* is a mental representation, the result of a perception process that entangles a relationship between the data and an internal representation model.

of them.Shütze&Sprouse (2013: 30) establish five major aspects in which typical informal linguistic judgment gathering differs from the standards of experimental cognitive science:

> It typically involves (i) relatively few speakers (fewer than ten), (ii) linguists themselves as the participants, (iii) relatively impoverished response options (such as just "acceptable," "unacceptable," and perhaps "marginal"), (iv) relatively few tokens of the structures of interest, and (v) relatively unsystematic data analysis.

Despite the highlighted differences, the authors claim,"it is not obvious what the 'best' choice is in each case" and argue that "all methods appear to provide relatively reliable results."

Cullicover and Jackendoff (2010), in turn, agree with the claim that subjective judgments are vulnerable to investigator bias, and that other types of data are relevant for linguistic analysis. Nevertheless, they claim that informal judgments are essential in formulating linguistic theories and are often sufficient for theory development. The data obtained by means of grammaticality judgments are, thus, the raw material for hypotheses about the structure of the language faculty. According to Cullicover and Jackendoff, without such judgments, the experimental enterprise cannot get off the ground.

2. JUDGMENTS AND DECISION-MAKING: THE PSYCHOLOGICAL SIDE

The judgment and decision-making field is a research area in behavioral science that focuses on how people (and also other organisms and/or machines) combine desires (utilities, personal values, goals, ends) and beliefs (expectations, knowledge, means) to choose a course of action (Hastie 2001:655-6). Conceptually, making a decision relies on three components: (i) courses of action, including choice options and alternatives; (ii) beliefs about objective states, processes, and events in the world; and (iii) desires, values, or utilities that describe the consequences associated with the outcomes of each action-event combination.

Evans (2008) highlighted that dual-processing accounts of human behavior are very common in cognitive and social psychology literature about higher cognitive processes, such as thinking, reasoning, decision-making, and social judgment. Dual accounts establish a distinction between two kinds of cognitive processes: processes that are automatic, fast, and unconscious and processes that are non-automatic, slow, deliberative, and conscious.Different judgment tasks may recruit distinct cognitive operations and the process of decision-making—in which both the judgment and the choice between the available alternatives are at stake—may be influenced by several factors.

Regarding acceptability judgments, an important question is to what extent they rely on true "speakers' intuitions"—triggering, therefore, unconscious and automatic processes—or, on the contrary, activate more deliberative and conscious mechanisms. As pointed out by Schütze and Sprouse (2013), the fact that acceptability judgments demand metalinguistic awareness does not undermine their external validity. However, acareful interpretation of the obtained results is necessary, particularly when there is no strict control of the stimuli, a small number of participants, no statistical analysis, etc.

It is worthwhile to note that there are relevant differences regarding the type of acceptability judgment used. We can divide judgment tasks in two major categories: (i) non-numerical tasks, such as the force-choiceand yes-no task, and (ii) numerical tasks, such as *Likert*[6] scaling and magnitude estimation. While non-numerical tasks are designed to detect qualitative differences between conditions, numerical tasks provide information about the size of the difference observed.The choice of the judgment task is directly related to the research question and there is no *a priori* best task. However, it is necessary to be cautious about the particularities of each type of task. For instance, in the case of the force-choice task, the fact that more than one option/sentence is simultaneously presented requires that the participant maintaintwo or more representations/structures in the memory and activate an inhibitory process when choosing the best option. In *Likert* scale rating, in turn, factors such as scale compression (using only intermediary points of the scale, for example) can obscure the results.

Another crucial point to be taken into account concerns the nature of the linguistic phenomenon that is being investigated. There is an obvious difference between assessing whether a sentence like (1)could or could not be produced by a native BrP speaker and evaluating more fine nuances, like those in the example in (2), in which the acceptability of the second sentence depends on the possibility of ascribing a generic reading to the quantifier (examples extracted from Pires de Oliveira 2003a).

(1) *Criança chora toda.
    'Any child cries.'

(2) Toda criança chora. ??Ela / *Elas sente(m) medo.
    'Any child cries. She/They feel fear.'

---

[6] A *Likert* scale is a psychometric scale that is commonly involved in research that employs questionnaires.

Judgments concerning semantic factors are particularly problematic.Matthewson (2004: 370) claimed that: "semantic facts are often subtle, are usually context-dependent, and are almost never accessible by direct native-speaker intuitions (i.e., one cannot simply ask questions of the form 'What does X mean?')."In such cases, the researcher cannot simply present a sentence and ask the participant to decide if it is or it is not acceptable. Accessing the speaker's intuition may thus require a less direct type of judgment. An experimental approach can be a good option in these cases, since it provides a means for capturing more automatic responses and for controlling the influence of non-relevant variables.

3. JUDGMENT TASKS IN LINGUISTIC AND PSYCHOLINGUISTIC RESEARCH: Q-EXPRESSIONS

The nature of quantified expressions (from here, Q-expressions)—formed by universal and/or indefinite Qs, such as *all*, *every*, *each*, and *a*—is a topic in the syntax-semantics interface that has been investigated from the perspectives of both linguistic theory and sentence processing. Q-interpretation has been explored by a number of psycholinguistic studies, mainly in language acquisition (Phillip 1995; Crain et al.1996; Roeper, Strauss, and Zurer Pearson 2006; among several others). Several studies suggest that children have apparently non-adult readings for Qs. The "Q spreading" phenomenon has been given syntactic, semantic, and pragmatic explanations (see Roeper, Zurer Pearson, and Grace 2011). Nevertheless, despite the evidence related to the acquisition of these items, Q processing does not seem to be trivial even for adults. In fact, there is also evidence of "childlike behavior" by adults interpreting universal-Qs (Brooks and Sekerina 2006). Q processing, therefore, seems to pose some level of difficulty for both children and adults.

Traditionally, Qs are classified as universal or indefinite. Universal-Qs denote the totality of applicable values associated with the expression they take as a complement. These elements show a syntactic distribution that partially overlaps with definite Determinants (D), as the two items share the same definiteness constraints. Therefore, universal-Qs are classified as strong, as opposed to weak or indefinite-Qs. Universal-Qs belong to a limited paradigm. In BrP the following are universal-Qs: *todo*(s) 'every'/ 'all',*cada* 'each',*ambos* 'both', and *nenhum* 'no'/ 'none'.

The interpretation of Q-expressions, as well as the scope preferences—in the case of sentences that contain more than one Q such as***Every kid climbed a tree***—have been

investigated in the context of both linguistic and psycholinguistictheory. Intuitive judgments have been employed in several studies of scope preferences among adults and children, while other research has used more controlled experimental settings. In the next sections we present: (i) the descriptions of the three Qdiscussed here—*cada, todo*and *todos os*—available in the theoretical literature, and (ii) a set of experimental results—related to the interpretation of those expressions—derived from studies conducted in the fields of experimental syntax/semantics and psycholinguistics, respectively.

3.1. Q-EXPRESSIONS IN BRP: INFORMAL LINGUISTIC JUDGMENTS

In order to enlighten the differences between informal judgments and those obtained from the use of experimental—more controlled—methodology, in this paper we discuss the case of universal BrP Qs.For such, we consider three of the universal BrP quantifiers, namely:

(i) *cada*, which can be translated as the English Q'each';

(ii)*todo* + singular NP, which is roughly equivalent to the English Q'every', although—according to Müller, Negrão, and Quadros Gomes (2007)—they are not synonymous, and;

(iii) *todo* + plural DP, which has the approximate meaning of the English Q'all'.

Although there is agreement in the literature regarding the quantifier *cada*—characterized as a true distributive Q (Negrão 2002; Quadros Gomes 2009; Lima 2013)—the nature of the universal-Q *todo* and its plural version *todos osx* is a subject of theoretical debate (Negrão 2002; Pires de Oliveira 2003a, 2003b; Müller, Negrão, and Quadros Gomes 2007; Quadros Gomes 2009; Lima 2013).

Müller, Negrão, and Quadros Gomes (2007) consider that, in BrP, the same morphological item*todo* can bear different features and combine with either a singular NP or a singular or plural DP as its argument, yielding different interpretations. The authors defend that in all the contexts, *todo* is a universal distributive Q[7].

(3) Toda família construiu uma jangada.(*Todo* + singular NP)

Q-singularfamily-singular built a raft.

'Every family built a raft.'

---

[7] It is worth mentioning that the author's examples contain the collective noun "family." The "always distributive reading"in these examples depends crucially on the fact that there is more than one member in the family, so the following interpretation is also possible: "each of the members of x built a raft."

(4) Todas as famílias construíram uma jangada. (*Todo* + plural DP)

Q-~plural~ the-~plural~ family-~plural~ built a raft.

'All the families built a raft.'

(5) Toda a família construiu uma jangada.(*Todo* + singular DP)

Q-~singular~ the-~singular~ family-~singular~ built a raft.

'All of the family built a raft.'

In this perspective, *todo* is a distributive Q that performs the same operation in all contexts in which it occurs. It is able to quantify over parts of both its restriction and its nuclear scope, and it relates these two parts (# from a Q like *every* or *each*, or from *cada* in BrP). The different readings of the sentences it participates in stem from the different denotations of its arguments. In this analysis, *todo*, when combined with a singular NP, never gets a collective reading because the NP is a predicate and does not denote an entity that could be involved in a collective action.This approach has the advantage of not posing any ambiguity for the quantifier *todo*, and of compositionally deriving its effects in distinct contexts.

According to Negrão (2002), *cada* is a distributive universal Q, while *todo* exhibits a chameleonic behavior, licensing both a distributive (6) and a generic reading (7). Negrão claims that Q-expressions with *todo* in BrP behave as indefinites (in the sense of Heim 1982).

(6) Todo aluno leu um texto (= *Cada aluno leu um texto*)

'Each student read a book.'

(7) Todo homem é inteligente.

'Everymanisintelligent.'

Pires de Oliveira (2003a) characterizes*todo*and*todos os* as distincttypesof universal-Qs. She argues that these Qs behave differently regarding flotation—compare examples (8) vs. (9)—and anaphoric recovering, see sentences (10) and (11).

(8) *Criança toda chora.

'Any child cries.'

(9) As crianças todas choram.

   'The children all cry.'

(10) Toda criança chora. ??Ela/ *Elas sente/m medo.

   'Any child cries. She/They feel fear.'

(11) Todas as criança choram. Elas sentem medo.

   'All the children cry. They feel afraid.'

Pires de Oliveira does notconsider—*contra* Negrão 2002—*todo* anindefinite. For Pires, *todo* plus NP is not licensed in episodic sentences, and it cannot occupy the object position (12):

(12) a. *Toda criança se machucou.

   'Every child got hurt.'

   b. * Ele canta toda canção.

    'He sings every song.'

When *todo*plus NP is modified by a partitive in episodic sentences, the sentence acceptability gets better (13), and when the modifier is a relative, *todo* becomes acceptable in existential contexts (14). The same occurs with *todo* plus NP in object position (15):

(13) ? Toda criança da festa se machucou.

   'Every child atthe party got hurt.'

(14)a. *Toda criança está brincando.

   'Everychildisplaying.'

   b. Toda criança que veio para a festa está brincando.

   'Every child who came to the party is playing.'

(15) a. *João conversou com toda mulher.

    'John talked to every woman.'

    b. João conversou com toda mulher que ele encontrou.

    'John talked to every woman who came up to him.'

In both cases—episodic sentences and object position—the expression formed by *todo* plus NP is interpreted as some sort of temporalized "universal." Pires de Oliveira isolates the

specificity from quantification. According to her, *todo* could belong to a class of non-specific (perhaps modal) Qs.

Quadros Gomes (2009), in turn, considers that *todo* is not a canonical operator like *cada*, but a modifier of the predicate relation. While English 'every' andBrP*cada* would establish distributive relations between two phrases,*todo* would be a modifier of an existent relation.

Taken as a whole, the different proposals suggest that there is no consensus in the theoretical characterization of the universal-Q *todo* and its plural version *todos*. In the next section, we present experimental results that can contribute to understanding the nature of theseBrP Q-expressions.

3.2. Q-EXPRESSIONS IN BRP: EXPERIMENTAL DATA

The processing of quantified sentences, particularly scope interaction effects, has been widely investigated in English, both in research conducted with adults (Ioup 1975a, 1975b; VanLehn 1978; Gil 1982; Gillen 1991; Kurtzman and MacDonald 1993; Tunstall 1998; Kemtes and Kemper 1999; Anderson 2004; Dwivedi et al. 2010; Dwivedi 2013; Brooks and Sekerina 2006; among others) and in the language acquisition literature (Philip 1995; Crain et al. 1996; Roeper, Strauss, and Zurer Pearson 2006; Brooks and Sekerina 2006; among others). In BrP, however, as far as we know, apart from our own research (Marcilese and Rodrigues 2014a, 2014b; Rodrigues and Marcilese 2014b, submitted; among others), there are only two studies based on experimental methodology (Lopes 2013;Lima 2013), both of them within the field of experimental syntax/semantics.

The experimental data reported in this section were collected from four major types of experimental tasks and all of them involve some kind of linguistic judgment. We synthesize the results according to the tasks used in the data collection:

(a)   *Sentence or picture selection tasks*: These are tasks of the "forced choice" type, because the participants must choose an answer between a set of options previously presented;

(b)   *Sentence-picture comparison tasks,*also named sentence "verification tasks" (Clark and Chase1972): In these, the participants should assess whether a given picture is consistent with the information presented linguistically. In this case, the subject must decide whether or not there is an identity relationship between two stimuli of different natures (visual and linguistic);

(c)   *Continuation sentence acceptability judgments,* also called *continuation consistency task* (Kurtzman and MacDonald 1993): This kind of task requires the prior presentation of a set of information followed by a target sentence. Participants have to judge to what extent the target sentence is a good continuation taking into account the information initially received; and; finally,

(d)   *(Plain) Sentence judgment task*: This is the most traditional judgment task, in which a single linguistic stimulus is presented and the participant must decide whether it could or could not be produced by a native speaker of that language. Both numerical and non-numerical responses could be employed (Yes/No judgments and scales).

Next, we briefly present the main experimental findings of the available studies. For further details of each study, see Table 1, in which we summarize the major information from each paper.

Lopes (2013) investigated the interpretation of sentences with the universal-Q *todo* plus singular NP and plural DP through an offline written questionnaire in a *sentence selection task* of type (a). In this study, participants had two possible paraphrases for each sentence and they could choose one or both of them (16); no pictures or additional information was presented:

(16) Toda menina está em um barquinho/Todas as meninas estão num barquinho.

Q-singular. girl-singular is in a boatQ-plural the-pluralgirls-plural. are in a boat.

'Every girl is in a boat' **OR** 'All the girls are in a boat.'

Há quatro meninas e todas elas estão num mesmo barquinho (   )

'There are four girls and all of them are in the same boat.'

Há quatro meninas e cada uma delas está em um barquinho distinto (   )

'There are four girls and each of them is in a different boat.'

The results are compatible with the pattern predicted by the theoretical description provided by Müller, Negrão, and Quadros Gomes (2007): *todo* plus singular NP favors distributive readings while *todo* plus plural DP promotes mostly collective readings. For *todo*, the results pointto a preference for the distributive reading (97.5% distributive and 15% collective). For *todos os*, the results suggest a preference for the collective reading (47.5% distributive vs. 97.5% collective). The quantifier *cada* was not evaluated in Lopes' (2013)

study, fillers were not used, and the results were not submitted to statistical analysis. Moreover, in the alternatives presented to the participants, Lopes used structures similar to the ones tested.

Marcilese and Rodrigues (2014b) conducted an equivalent experiment, but with other type of stimuli—pictures instead of sentences—to verify the interpretation of the same universal-Q. The visual options presented to the participants consisted of three different pictures(distributive/collective/filler). Different from Lopes' study, only one picture could be chosen.Belowwepresenta sampleof the material:

(17) Toda chave está numa garrafa **OR** Todas as chaves estão numa garrafa.

‘Every key is in a bottle’/ ‘All the keys are in a bottle.’
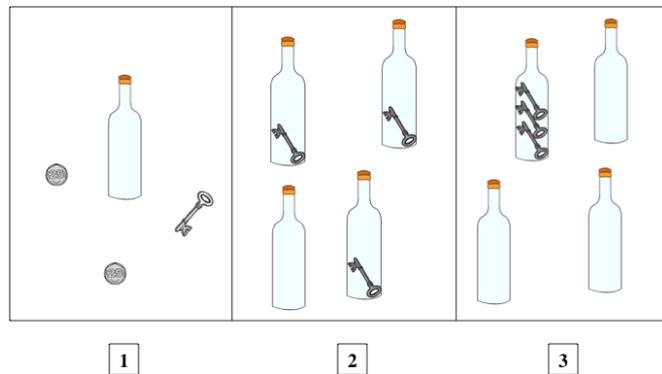
There was a clear preference for collective images for both *todo*+ plural DP and *todo*+singular NP sentences. Notice that, although more than one alternative was provided (as in Lopes' study), participants could choose only one option, and there was pressure to answer as fast as possible. Furthermore, we tried to access the participants' interpretation more indirectly, without using linguistic sentences that contain the same type of structure under investigation. In a follow-up experiment (Marcilese and Rodrigues 2014b), we conducted a *sentence-picture comparison ta*sk, with the same type of visual stimuli used in the picture selection task. Different from the previous task, in order to reduce the cognitive demand related to the simultaneous presentation of images, participants saw only one picture in each trial.

(18)
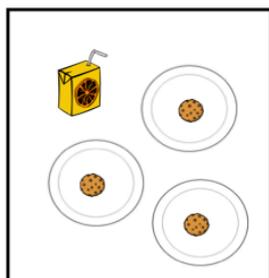
Distributive picture
Todo biscoito estava num prato **OR**
Todos os biscoitos estavam num prato
'Every cookie was in a plate'.
'All the cookies were in a plate'.
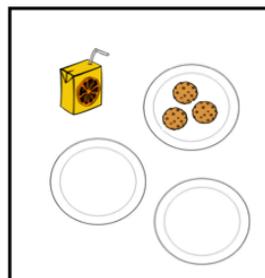
Collective picture
Todo biscoito estava num prato **OR**
Todos os biscoitos estavam num prato
'Every cookie was in a plate'.
'All the cookies were in a plate'.

As a whole, the results were compatible with what was obtained in the picture selection experiment. For the Q-expression *todosos*, there was a strong preference for the collective choice: 92% of YES judgments for this configuration compared to 46% for the combination with distributive pictures. For *todo* + NP, the mean of YES judgments for the collective picture was also quite high (89%) when compared to 56% for the distributive one. Regarding reaction time, there were significant differences between the distributive and collective pictures, both for the plural and singular Q-expressions. The reaction time for YES judgments was significantly higher for distributive pictures in general, suggesting that this was the less accepted option.

Lima (2013) investigates the semantic properties of *cada* ('each') and *todos os x* ('all the x') by means of a written questionnaire employing*continuation sentence acceptability judgments*. A sample of the stimuli is in (19) below:

(19)

Joint context [Collective]
As crianças tiveram uma atividade extra hoje na escola. Elas tinham que participar da construção de uma jangada parecida com uma jangada indígena. As crianças trabalharam na mesma jangada: Maria cortou a madeira, João colocou a parte lateral, Pedro colou a parte inferior e assim foi... até uma jangada ficar pronta.

[The children had an extra activity today at school. They had to participate in the construction of a raft similar to an indigenous raft. The children worked on the same raft: Maria cut the wood, João put together the lateral part, Pedro put together the lower part… until a raft was ready].

Separatelycontext [Distributive]

As crianças tiveram uma atividade extra hoje na escola. Elas tinham que construir individualmente uma jangada parecida com uma jangada indígena. Emumaclasse com 30 alunos, 30 jangadas foram construídas.

[The children had an extra activity today in the school. They had to build individually a raft similar to an indigenous raft. In a class with 30 students, 30 rafts were built.]

Sentence-trial options for both contexts:
( ) Cada criança construiu uma jangada ('Eachchildbuilt a raft').
( ) Todas as crianças construíram uma jangada ('All the childrenbuilt a raft').

Lima's results reveal clear differences between *cada* and *todosos x* inserted in distributive and collective contexts: Q-*cada* was the preferred option in distributive contexts (46% were instances of *cada* only) and this Q was never chosen in collective contexts. In contrast, *todos os* was the only choice in collective contexts (100%), and it was the second choice in distributive contexts (but never as the only option). Lima (2013) states that *cada* is a Q marked essentially for distributivity, and that the root *tod-* is not a Q, but a modifier that inherits collective or distributive readings from the plural DP with which it is combined. Lima tested a small number of participants (just eight); she did not explore Q*todo*+singular NPand did not provide a statistical analysis of the data.

Rodrigues and Marcilese (2014b, submitted)also conducted a *continuation sentence acceptability judgment task*, inspired by the work of Lima (2013), but in this case, we investigated the processing of three Q-expressions (*cada*, *todo*, and *todos os*) among a larger number of participants and with a slightly different methodology. Participants had to judge the suitability of a continuation sentencewith one of the investigated Q-expressions. The context and the type of Q-expression were within-subject factors. Below, we provide a sample illustrating the two discursive contexts:

(21)

CollectiveContext
Maria tem um pequeno restaurante. Ela resolveu enfeitar o balcão onde fica o bar. Em cima do balcão, colocou um vaso de porcelana com lindas flores. Os frequentadores do restaurante adoraram a decoração quando chegaram para o jantar.
Sentence-trialoptions: *Cada flor estava em um vaso de porcelana/ Toda flor estava em um vaso de porcelana/ Todas as flores estavam em um vaso de porcelana.*

['Mary owns a small restaurant. She decided to decorate the bar and put there a porcelain vase with beautiful flowers. The customers loved the decor when they arrived for dinner.'Sentence-trial options: 'Each/Every/All (of the) flower

(singular/plural) was/were in a porcelain vase.']

DistributiveContext

Pedro trabalha numa lavanderia e se ocupa de organizar as roupas limpas e passadas. Hoje ele preparou o pedido de um hotel que entrega as toalhas de banho para seus hóspedes em saquinhos plásticos individuais personalizados. Pedro dobrou e embalou as toalhas do hotel.
Sentence-trialoptions: *Cada toalha estava num saquinho plástico/ Toda toalha estava num saquinho plástico/ Todas as toalhas estavam num saquinho plástico.*

['Peter works in a laundry and he's responsible for the clean clothes. Today, he organized the order from a hotel that delivers the bath towels for guests in personalized individual plastic bags. Peter folded and packed the hotel towels.' Sentence-trial options: 'Each/Every/All (of the) towel/s was/were in a plastic bag.']

Different from Lima's (2013) study, the participants did not have to compare possibilities, since only one type of Q was presented after the discursive context. We tested an expressive number of participants (forty-two),and we also measured the response times. As expected, there was a clear preference for *cada* in distributive contexts. For the expressions with the Q *todo*, a marginal difference was found between the number of positive judgments for collective and distributive contexts in only the one-tail analysis (p=0.042). For the *todos os* expression, there was no difference between the number of positive and negative responses for either collective or distributive contexts; that is, participants accepted thisQ indistinctively after distributive or collective readings.Reaction times are more informative, particularly in the case of *todo* + singular NP. We verified that participants significantly took more time to reject the Q-expression after collective than after distributive contexts, which points to a preference—in terms of time measure—for collective contexts. In the case of *todo* + plural DP, we didnot observe differences between the conditions, as if the two possibilities (distributive and collective) were equally acceptable. This experiment reveals the relevance of analyzing not only response times, but also response timesin relation to the type of response given (positive/negative).

Complementarily, we also conducted an offline written questionnaire in which the participants (forty-four adult speakers) had to rank the provided options. In this case, the three Q-expressionswere presented at the same time. The context was available for the entire duration of the task, including the time at which the sentence was to be ranked. Participants were permitted to select more than one option for the same ranking. For distributive contexts, *cada* was most often selected as the best option, and this preference was significant, withtodos os in second place. For collective contexts, the most preferred option wastodos os. The quantifier

todo exhibited no clear pattern of behavior, with a similar distribution among first, second, and third place rankings in distributive contexts, and predominantly second place rankings in collective contexts.

In Marcilese and Rodrigues (2014a), we report the results of a non-cumulative word-by-word self-paced reading experiment combined with a sentence judgment task. The participants were presented with conjoined sentences: The first sentence contained the Q-expression and the second one, a singular or plural anaphoric pronoun, as in the sample below.

(22) Todas as bailarinas estavam num camarim e **ele(s)** foi/foram esvaziado(s) pela coreógrafa.

'All the dancers were in a dressing room and it/they was/were emptied by the choreographer.'

Participants read the sentences word by word using the space bar to go forward. After the presentation of the conjoined sentences, a happy (green) and a sad (red) face emoticon appeared on the screen and the participants had to press keys with equivalent colors on the keyboard. In this experiment, we analyzed both reading and response times and the type of judgment. The independent variables were *type of Q* (*cada*, *todo* + singular NP, and *todo* + plural DP) and *number of the pronoun* (singular/plural). Anaphoric retrieval was taken as an index of the interpretation attributed to the scope relations established in the first sentence—the singular form of the pronoun was related to a joint/collective reading, and the plural form was interpreted as corresponding to a distributive reading.

Concerning the self-paced reading task, we analyzed the reading times at three different points: at the anaphoric pronoun, at the auxiliary, and also at the participle. The statistical analysis revealed significant effects of the independent variables (for further details, see Marcilese and Rodrigues 2014a), namely higher reading times for plural pronouns (= distributive reading), regardless of the Q configuration and faster reading times for singular pronouns (= collective readings) in the *todo* + plural DP condition. Judgments also revealed differences between the Qs. For *cada*, a preference was verified for distributive readings (82%), although the number of positive judgments was also above the level of chance for collective readings (an unexpected result). No difference between the conditions was obtained for judgment time. For *todo* + singular NP, no clear pattern was observed: although the number of positive judgments was higher for singular pronouns (collective = 47.5%), the percentage was above 50%. Higher mean reading times for distributive readings is in consonance with the lower degree of acceptability of this option (19%).In the case of *todo* + plural DP, there was a clear

preference for collective readings—91.5% positive judgments in the singular pronoun condition (vs. 25% positive judgments in the plural condition). Higher mean reading times for distributive readings are compatible with the judgments results (a low degree of acceptability of distributive interpretations).

We also compared these results with those of an offline written questionnaire with an ordinal 1–5 scale (1= *bad* and 5= *excellent*), in which we used the same type of sentences tested in the self-paced reading experiment. There was a clear preference for distributive readings in the case of *cada*: a greater proportion of judgments of types 4 and 5 (= *good* and *excellent*) for plural pronouns. For *todo* + plural DP, judgments also suggested a preference, in this case, for collective readings: a greater proportion of judgments 4 and 5 for singular pronouns. *Todo* + singular NP, in turn, did not exhibit a clear pattern, a similar result to what we verified in the online task. Notice that in both the online and offline tasks, we tried to capture the reading preferences more indirectly: Instead of providing a possible interpretation—by means of a sentence (Lopes 2013) or visual stimuli (Marcilese and Rodrigues 2014b)—we provided continuations for the sentences and tested the acceptability of the whole structure. The rationale was that the anaphoric recovery could reactivate a representation of an antecedent generated from the computation of sentences with QPs.

Study	Investigated Q-expression		
Sentence or picture selection tasks	*cada*	*todo* + sg. NP	*todo* + pl. DP
Lopes (2013) – Sentence selection Offline written questionnaire with two sentence options (distributive/collective); participants could choose one or both of them. Four trials in two conditions (singular and plural) and no fillers. Number of participants: 20. No statistical analyses.	Not evaluated.	Preference for distributive reading (97.5% distributive; 15% collective).	Preference for collective reading (47.5% distributive; 97.5% collective).
Marcilese and Rodrigues (2014b) – Picture Selection Online task[8] (reaction times were recorded) with three picture options (distributive/collective/filler); only one picture could be chosen. 12 trials (6per condition) and 36 fillers.	Not evaluated.	Preference for collective pictures (mean of target responses = 4.05; Max Score = 6).	Preference for collective pictures (mean of target

[8] We employ the term "online task" in two different senses here: (i) in contrast to "offline" questionnaire studies in which the participants have total control of time and task execution rate; and (ii) in the traditional sense in cognitive science (e.g., in the case of self-paced reading tasks).

	cada	todo + sg. NP	todo + pl. DP
Number of participants: 20 Statistical analyses: all the reported results are statistically significant.			responses = 5; Max Score = 6).
Sentence-picture comparison task	*cada*	*todo* + sg. NP	*todo* + pl. DP
Marcilese and Rodrigues (2014b) Online task (reaction times were recorded) 20 trials and 60 fillers. Number of participants: 20 Statistical analyses: significant differences in both, online and offline measures.	Not evaluated.	Preference for collective pictures (89% vs. 56% for distributive pictures).	Strong preference for the collective pictures (92% vs. 46% for distributive pictures).
Continuation sentence acceptability judgments	*cada*	*todo* + sg.NP	*todo* + pl. DP
Lima (2013) Offline written questionnaire Two different discursive context (joint = favoring a collective reading and separately = favoring a distributive reading) and two possible continuations; participants could choose one or both of them. Number of participants: 8 8 trials and 16 fillers *per* list. No statistical analyses.	Preferential option in distributive contexts(46%). Never chosen in collective contexts.	Not evaluated.	Only choice in collective context (100%), and 2nd option in distributive context (never considered the only option).
Rodrigues and Marcilese (2014b, submitted) Online task (reaction times were recorded) Number of participants: 42 10 context/sentence trials and 20 fillers. Statistical analyses: all the reported results are statistically significant.	Preference for distributive contexts.	Marginal difference between the number of positive judgments for collective and distributive contexts.	No difference between collective and distributive contexts.
Offline written questionnaire with sentence ranking Number of participants: 44 10 context/sentence trials and no fillers. Statistical analyses: all the reported results are statistically significant.	1st option in distributive contexts.	No clear pattern. Similar distribution between 1st, 2nd, and 3rd position in distributive contexts; 2nd option in collective contexts.	1st option in collective contexts; 2nd option in distributive contexts.
(Plain) Sentencejudgmenttask	*cada*	*todo* + sg.NP	*todo* + pl. DP

| Marcilese and Rodrigues (2014a)

Online task (reaction times were recorded): Non-cumulative word-by-word self-paced reading.
Number of participants: 50.
16 trials and 32 fillers.
Statistical analyses: all the reported results are statistically significant. | Reaction time: there was no difference between the conditions.

Mean of positive judgments: preference for distributive reading(82%); collective reading = 67.75%. | Reaction time: Higher mean reading times for distributive readings (lower acceptance of this option).

Mean of positive judgments: no clear pattern (collective = 47.5%, distributive reading = 19%). | Reaction time: Higher mean reading times for distributive readings (lower acceptance of this option).

Mean of positive judgments: preference for collective reading (91.5%); distributive reading = 25%. |
|---|---|---|---|
| Offline written questionnaire with an ordinal 1-5 scale (1= *bad* and 5= *excellent*)
Number of participants: 51
8 trials and 32 fillers.
Statistical analyses: all the reported results are statistically significant. | Greater proportion of judgments of types 4 and 5 (= *good* and *excellent*) for distributive reading. | No clear preference towards a distributive or a collective reading. | Greater proportion of judgments 4 and 5 for collective reading. |

Table 1: Summary of the main experimental results.

The results reported in this section seem to support the idea that each Q has an effect on scope preferences. In the case of *cada*, a clear preference for a distributive reading was observed, a result that has been replicated in several kinds of experimental tasks. *Cada* may thus be defined as a true distributive-Q.*Todos os*, in turn, was compatible with both collective and distributive readings when inserted in the respective scenarios (Lima 2013; Rodrigues and Marcilese, submitted). This picture arises only when an explicit context—collective or distributive—is provided. In general, we have observed a strong preference for collective (or non-distributive) readings, and we hypothesized that doubly quantified sentences could be subject to shallow processing, depending on the task. In such a case, participants would interpret the expression *todos os x* on the basis of the lexical-semantic properties of the Q.

Turning to the Q-expression *todo* + singular NP, this did not exhibit a clear pattern. Although a distributive reading can be associated to this Q (as verified in Lopes'(2013) offline

questionnaire test), this is not the only possible reading. In our sentence picture selection and sentence-picture comparison task, we verified a high number of responses in which a collective/non-distributive picture was associated with *todo* + singular NP. In the online self-paced reading task, participants do not clearly show a preference for collective or distributive readings and, in the offline written questionnaire, *todo*+NP is not chosen as the first option when it is contrasted with *cada* (the preferred option in the distributive condition) or *todo*+plural DP(the first choice option in the collective condition).

Universal-Qs such as *each* and *every* are not distinguished in all languages—for example, in Hebrew they are not—but in those that have two different items for these Q, they exhibit different semantic and lexical properties (Novogrodsky, Roeper, and Yamakoshi 2013). In English, *each* is a distributive Q in all contexts, while *every* allows—as well as *all*—both distributive and collective readings. In other words, *every* "accommodates" collective interpretations (22a), although it doesnot easily accept collective predicates (22b).

(22) a. The mom took *every* dog for a walk.

   Distributive interpretation: one by one.

   Collective interpretation: together.

   b. *Every student gathered in the hall.

   c. *All* the students gathered in the hall.

Our experimental findings concerning the BrP Q-*todo* seem to be compatible with what was observed with *every*in English. As a common feature, both Qs—*cada* and*todo* + NP in BrP, and *every* and *each* in English—require individuation, even when undergoing a collective interpretation. One additional aspect is the possibility to ascribe a generic reading to *todo*, as observed by Negrão (2002).

4. FINAL REMARKS

Universal-Q processing is not a trivial task for either children or adults. Experimental situations, due to time pressure (in online studies) and/or to the nature and type of stimuli (e.g., visual or linguistic), can bring additional processing demands and make Q interpretation even more costly for participants. For the Q-expressions investigated here, the amount of information

associated with the Q would certainly be different among the several tasks employed in each study.

The main differences between the reported studies are: the number of participants and experimental items, the use of fillers, the selected dependent variables (mean of target-judgments, mean of judgment time, type of judgment, etc.), and the statistical processing of the data. Another relevant difference has to do with the dynamics of the task itself. On the one hand, certain tasks entangle slower, deliberative, and conscious processes. On the other hand, there are tasks that aim to access more automatic and unconscious processes. This second type also provides measures more directly related to linguistic processing, such as reading time and judgment reaction time. Thereby, when defining the type of task and analyzing the results, these aspects should be taken into account and, depending on the nature and suitability of the linguistic phenomenon, it is productive to combine controlled offline judgments and online tasks.

In addition to the specificity of the task itself, another aspect that may be relevant in the investigation of certain linguistic phenomena relates to the characteristics of the experimental group. Age-related differences appear to be highly relevant in the evaluation of certain linguistic aspects. Specifically with respect to the case analyzed in this paper, besides the already mentioned non-adult comprehension pattern found in children (Philip 1995; Roeper, Strauss, and Zurer Pearson 2006; Brooks and Sekerina 2006; among others), differences in Q-expression processing among adults have also been identified as a function of age. Kemtes and Kemper (1999) compare young and older adults' processing of complex sentences involving Q scope ambiguities (younger sample: age range from 18 to 25; older sample: age range from 65 to 85 years) by means of a *continuation sentence acceptability judgments task* (e.g.,*Every actor used a prop… The prop(s) was/were on the stage* OR *An actor used every prop… The prop(s) was/were on the stage*). The results reveal similar patterns of reading times, but differences in the preferred continuation: young adults preferred continuations postulating multiple entities (a distributive interpretation), whereas older adults preferred continuations with a single entity (suggesting a collective interpretation). These results seem to be compatible with the findings reported by Novogrodsky, Roeper, and Yamakoshi (2013), in this case with children: at the age of 6–7 years, children prefer the collective interpretation to the distributive interpretation for *every,* while young adults accept both the collective and distributive interpretations. Working memory and executive control mechanisms are factors directly influenced by age, and they have a crucial role in language processing (Rodrigues 2011; Rodrigues and Marcilese 2014a).

Together, the results of the experiments reportedherereinforce the relevance of analyzing the same phenomenon by means of different techniques. Experiments can provide evidence concerning constructions about which speakers (even though linguists) do not have consistent informal judgments. Nevertheless, even when all of the standard experimental controls are taken into account, different experimental design and procedures can recruit different cognitive processes (more or less algorithmic) and, as a consequence, different patterns can emerge.

REFERENCES

1. ANDERSON, C. The Structure and Real-time Comprehension of Quantifier Scope Ambiguity. Ph.D. Dissertation, Northwestern University, 2004.

2. BRESNAN, J. Is syntactic knowledge probabilistic? Experiments with the English dative alternation. In: FEATHERSTON, S. & STERNEFELD, W. (Eds.). Roots: Linguistics in Search of Its Evidential Base. Berlin: Mouton de Gruyter, 77–96, 2007.

3. BROOKS, P. J. &SEKERINA, I. A. Shortcuts to Quantifier Interpretation in Children and Adults. Language Acquisition13(3), 177–206, 2006.

4. CHOMSKY, N. Aspects of the theory of syntax. Cambridge, Mass: M. I. T. Press, 1965.

5. CLARK, H. & CHASE, W. G. On the Process of Comparing Sentences Against Pictures. Cognitive Psychology 3, 472–517, 1972.

6. CRAIN, S., THORTON, R., BOSTER, C., CONWAY, L., LILLO-MARTIN, D. &WOODAMS, E. Quantification without qualification. Language Acquisition3(2), 83–153, 1996.

7. CULICOVER, P. & JACKENDOFF, R. Quantitative methods alone are not enough: Response to Gibson and Fedorenko. Trends in Cognitive Sciences 14(6), 2010.

8. DWIVEDI, V. D. Interpreting quantifier scope ambiguity: Evidence of heuristic first, algorithmic second processing. PLOS One, v. 18, n. 11, 2013.

9. DWIVEDI, V. D., PHILLIPS, N. A., EINAGEL, S. & BAUM, S. R. The neural underpinnings of semantic ambiguity and anaphora. Brain Research v. 1311, p. 93–109, 2010.

10. EVANS, J. St. B. T. Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition. Annu. Rev. Psychol. 59:255–78, 2008.

11. GIBSON, E. & FEDORENKO, E. Weak quantitative standards in linguistics research. Trends in Cognitive Sciences 14(6), 233–234, 2010a.

12. GIBSON, E. & FEDORENKO, E. The need for quantitative methods in syntax and semantics research. Language and Cognitive Processes, 2010b.

13. GIBSON, E., PIANTADOSI, S. & FEDORENKO, E. Quantitative methods in syntax/semantics research: A response to Sprouse and Almeida (2012). Language and cognitive processes, 1–12, 2012.

14. GIL, D. Quantifier Scope and Linguistic Variation. Linguistics and Philosophy 5: 421– 472, 1982.

15. GILLEN, K. The Comprehension of Doubly Quantified Sentences. Ph.D. Dissertation, University of Durham, 1991.

16. HASTIE, R. Problems for judgment and decision making. Annu. Rev. Psychol. 52:653–83, 2001.

17. HEIM, I. The Semantics of Definite and Indefinite Noun Phrases. Dissertation, UMass Amherst, 1982.

18. IOUP, G. Some Universals for Quantifier Scope.In: KIMBALL, J. (Ed) Syntax and Semantics, 37-58. New York: Academic Press, 1975a.

19. IOUP, G. The Treatment of Quantifier Scope in a Transformational Grammar. Doctoral Dissertation, The City University of New York, NY, 1975b.

20. KEMTES, K. A. & KEMPER, S. Aging and Resolution of Quantifier Scope Effects. Journal of Gerontology: Psychological Sciences 54B(6), 350–360, 1999.

21. KURTZMAN H.S. & MacDONALD, M. C. Resolution of quantifier scope ambiguities. Cognition, 48: 243–279, 1993.

22. LEVELT, W. J. M. Some psychological aspects of linguistic data. LinguistischeBerichte 17, 18–30, 1972.

23. LIMA, S. (Maximality and Distributivity in Brazilian Portuguese. RevistaLinguística, 9(1), 133–157, 2013.

24. LOPES, R.E.V. Quantifiers in a nominal agreement language: acquisitional cues. In: STAVRAKAKI, S., KONSTANTINOPOULOU, P. &LALIOTI, M. (Eds.) Advances in Language Acquisition. Newcastle: Cambridge Scholar Publishing, 2013.

25. MARCILESE, M. & RODRIGUES, E. dos S. Correferência anafórica e interpretação de quantificadores universais. Cadernos de Letras da UFF, 49:109–131, 2014a.

26. MARCILESE, M. & RODRIGUES, E. dos S. The Interpretation of Brazilian Portuguese Quantifier 'todo' in Distributive and Collective Contexts: An Experimental Study. In: TORRENS, V. & ESCOBAR, L. (Org.). The Processing of Lexicon and Morphosyntax. 1ed.Newcastle upon Tyne: Cambridge Scholars Publishing, p. 69–88, 2014b.

27. MATTHEWSON, L.On the methodology of semantic fieldwork. IJAL, 70(4), 369–415, 2004.

28. MÜLLER, A. L.,NEGRÃO, E. V. &QUADROS GOMES, A. P. "Todo" em contextos coletivos e distributivos. DELTA 23(1), 71–95, 2007.

29. NEGRÃO, E. Distributividade e Genericidade nos Sintagmas Introduzidos por Cada e Todo. Revista do GEL, special number, 187–204, 2002.

30. NOVOGRODSKY, R., ROEPER, T. & YAMAKOSHI, K. The collective-distributive reading of each and every in language acquisition. In: STAVRAKI, S., LALIOTI, M. & KONSTANTINOPOULOU, P. (Org.). Advances in Language Acquisition. 1ed.Newcastle upon Tyne: Cambridge Scholars Publishing, 138–146, 2013.

31. PHILIP, W. Event quantification in the acquisition of universal quantification. Doctoral Dissertations, University of Massachusetts, Amherst, MA, 1995.

32. PHILLIPS, C. & WAGERS, M. Relating Structure and Time in Linguistics and Psycholinguistics. In: GASKELL, G. (Ed.). Oxford Handbook of Psycholinguistics, 739-756. Oxford University Press, 2007.

33. PIRES DE OIIVEIRA, R. Some remarksonTodo N in BrazilianPortuguese. Revista Letras 60, 363–384, 2003a.

34. PIRES DE OLIVEIRA, R. Istodo N in BrazilianPortuguese a quantifier? In: ANDERSEN, J., MENÉNDEZ-BENITO, P. & WERLE, A. (Eds.) The Proceedings of SULA 2, 99–116. GLSA, University of Massachusetts, Amherst, MA, 2003b.

35. QUADROS GOMES, A. P. O efeito grau máximo sobre os domínios: como todo modifica a relação argumento-predicado. Tese de Doutorado, Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2009.

36. RODRIGUES, E. dos S. O papel de mecanismos de controle executivo no processo linguistico: diferença de desempenho entre crianças e adultos em tarefas experimentais. RevistaLinguíStica 7(2):98–117, 2011.

37. RODRIGUES, E. dos S. &MARCILESE, M. Executive control in children's comprehension of ordinal modifier phrases. In: J. COSTA, A. FIÉIS, M.J. FREITAS, M. LOBO&A. L SANTOS. (Org.). New Directions in the Acquisition of Romance Languages: Selected Proceedings of the Romance Turn V. Newcastle upon Tyne: Cambridge Scholars Publishing, 2014a.

38. RODRIGUES, E. dos S. & MARCILESE, M. Procedimentos heurísticos no processamento linguístico? Um estudo sobre a interpretação de expressões quantificadas no PB. FórumLinguistico 11, 308–327, 2014b.

39. RODRIGUES, E. dos S. & MARCILESE, M. Contextual Information in Universal-Q Processing and Some Remarks on Distributivity, Collectivity, and Maximality, submitted paper.

40. ROEPER, T., STRAUSS, U. &ZURER PEARSON, B. The acquisition path of the determiner quantifier every: Two kinds of spreading. In T. Heizmann (Ed.) Papers in Language Acquisition. University of Massachusetts, Occasional Papers UMOP, 34. Amherst, MA: GLSA, 2006.

41. ROEPER, T., ZURER PEARSON, B. &GRACE, M. Quantifier Spreading is not Distributive. In:DANIS, N., MESH, K. &SUNG,H. (Eds.) Proceedings of the Boston University Conference on Language Development, 35, 526–539. Somerville, MA: Cascadilla Press, 2011.

42. SCHÜTZE, C. T. The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology. Chicago: University of Chicago Press, 1996.

43. SCHÜTZE, C. T. & SPROUSE J. Judgment data. InPODESVA, R. J. & DEVYANI, S. (Eds.), Research Methods in Linguistics. Cambridge University Press, 2013.

44. SPROUSE, J. & ALMEIDA, D. A quantitative defense of linguistic methodology, 2010. Available in: lingbuzz/001075

45. SPROUSE, J. & ALMEIDA, D. A formal experimental investigation of the empirical foundation of generative syntactic theory. Ms. University of California Irvine, 2011.

46. SPROUSE, J., & ALMEIDA, D. The empirical status of data in syntax: A reply to Gibson and Fedorenko. Language and Cognitive Processes, 2012a.

47. SPROUSE, J. & ALMEIDA, D. Power in acceptability judgment experiments. Unpublished Ms. University of California Irvine and New York University, Abu Dhabi, 2012b.

48. TREMBLAY, A. Theoretical and methodological perspectives on the use of grammaticality judgment tasks in linguistic theory. Second Language Studies, 24(1), 129–167, 2005.

49. TUNSTALL, S.L. The interpretation of quantifiers: semantics and processing. Ph.D. Dissertation, University of Massachusetts, Amherst, 1998.

50. VANLEHN, K. A. Determing the Scope of English Quantifiers. Technical Report, AI-TR- 483. Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA, 1978.

RESUMO: Tradicionalmente, teorias linguísticas de cunho formalista são baseadas em dados provenientes de julgamentos intuitivos do próprio linguista, relativosà boa-formação de sentenças numa dada língua. Julgamentos de gramaticalidade e aceitabilidade constituem um dos métodos de coleta de dados mais amplamente difundido e utilizado para verificar propostas teóricas. No entanto, os critérios de validade que subjazem a essa metodologia têm sido alvo de debate recente na literatura. O presente artigo visa a discutir esse tópico por meio da comparação entre propostas linguísticas – desenvolvidas principalmente a partir de julgamentos informais – e resultados experimentais, no que concerne à interpretação dos quantificadores universais *cada, todo* e *todos os* do português brasileiro. Resultados de um conjunto de experimentos sugerem que vários aspectos metodológicos – em especial, a presença ou ausência de pressões de ordem temporal impostas pelo tipo de tarefa experimental bem como demandas específicas da tarefa a ser executada – podem afetar o processamento, a interpretação e o julgamento de sentenças contendo sintagmas quantificados. No presente artigo, defendemos a ideia segundo a qual tanto linguistas quanto psicolinguistas exploram o mesmo domínio cognitivo, apesar de utilizar ferramentas distintas para atingir seus objetivos.

Palavras-chave: Aceitabilidade/Gramaticalidade; Julgamentos; Teoria Linguística; Psicolinguística; Quantificadores Universais; Português Brasileiro.