

WOULDN'T YOU AGREE WITH A BUNCH OF LEMONS? A SELF-PACED READING TEST ON AGREEMENT IN QUANTIFICATION EXPRESSIONS

Alessia Serafini
Francesca Franzon
Davide Bertocci
Giorgio Arcara
Chiara Zanini
Università di Padova

Abstract: *This paper investigates the potential involvement of semantics in the processing of agreement. Forty Italian speakers participated to a self-paced reading test. We examined the reaction times and acceptability ratings with respect to Italian quantification expressions such as 'a bag of / a bunch of'+ concrete/abstract noun. These expressions show two possibilities concerning Number agreement: the verb can agree either with the quantifier (syntactic agreement) or with the noun that it modifies (semantic agreement). When containing a concrete noun (e.g. bunch of lemons), these expressions are ambiguous between a reading in which the NP 'bag' is modified by a second NP denoting its content (i.e. a bag full of lemons) and a reading in which 'bag' quantifies over the NP head (i.e. lemons). The ambiguity disappears when the NP contains an abstract noun (e.g. a bunch of concepts). Results showed a preference for agreement with the concrete noun. We will discuss these findings arguing that there is no competition between semantic and syntactic information in the processing of agreement. Instead, the phenomenon can be explained in terms of syntactic ambiguity. We therefore conclude that agreement only takes place at the syntactic level.*

Keywords: *number agreement, syntactic-semantic interface, syntactic ambiguity*

1. Introduction

Current generative approaches define agreement as an operation taking place between two elements, namely a probe and a goal (Chomsky 2000; 2001). When the value of a feature of the probe changes, the correspondent feature of the goal changes accordingly. Agreement has generally been accepted by linguists as the exclusive dominion of the syntax (Pollock, 1989; Kayne, 1989; Chomsky, 2000); however, this claim is being challenged by psycholinguistic models that attribute agreement to other levels of the language, i.e. semantics.

In their work, Bates and colleagues (1982) claim, for instance, that different words in a given sentence may stand up as candidate-probes for agreement, due to some sort of competition among different agreement sources (for example syntax *vs* semantics). This proposal is also accepted by Badecker and Kuminiak (2007), who postulate in their Working Memory Retrieval Model a cue-based mechanism for retrieving the agreement source in the working memory.

In Haskell and MacDonald's (2003) proposal, called the Competition Model, the authors show that conflicts between semantic and grammatical information lead to increased variability in the verb form selection as well as longer processing times; they conclude that syntactic and non-syntactic factors interact in determining subject-verb agreement in their experimental sentences.

Yet, it is possible that instances of the alleged "semantic" agreement can be more properly accounted for in terms of a *syntactic ambiguity*. This ambiguity originates from the presence of multiple potential agreement controllers, i.e. two equally eligible NPs, and from a same linearization of two different possible syntactic structures. In particular, we will argue that the selection of the controller is not ruled by semantics, instead it depends on the underlying syntactic structure.

The goal of this study is therefore to address the question whether agreement is primarily a syntactic matter. We will try to solve the semantic/syntactic agreement issue by exploring the behaviour of Italian quantification expressions with *un mucchio di/ un sacco di* ('a bunch of/ a bag of') + NOUNPL, as in *un mucchio di limoni* ('a bunch of lemons').

It is well attested that in these expressions the verb can agree in number either with the first term of the quantification expression (1a) or with the second term (1b):

- | | | | | |
|--------|-----------------|------------------|---------------------------|--------------------|
| (1) a. | <i>Un sacco</i> | <i>di limoni</i> | <i>è stato venduto</i> | <i>al mercato</i> |
| | a bunch-SG | of lemon-PL | was sold-SG | at the marketplace |
| b. | <i>Un sacco</i> | <i>di limoni</i> | <i>sono stati venduti</i> | <i>al mercato</i> |
| | a bunch-SG | of lemon-PL | were sold-PL | at the marketplace |

There is a strong trend in performing agreement with the second member of a complex expression, as noted by Haskell et al. (2010) who described the Agreement Error Asymmetry. The author noted that participants produced more frequently agreement mistakes in the condition in which a verb wrongly agrees with a plural noun (as in **the key to the cabinets are missing*), than in the condition in which a verb incorrectly agrees with a singular noun (as in **the keys to the cabinet is missing*).

2. *The semantic/syntactic agreement issue*

Traditional grammars refer to agreement with the second term of Italian quantification expressions *un sacco di/un mucchio di* (conditions (1b) and (2b) in Table 1, below) as instances of a so called "semantic" agreement, and to agreement involving the first term (conditions (1a) and (2a) in Table 1) as instances of syntactic agreement.

Wouldn't You Agree With A Bunch Of Lemons? A Self-Paced Reading Test On
Agreement In Quantification Expressions

Condition	Verb agrees with:	Concreteness of noun	Examples
1a	First term	[+ concrete]	<i>Un sacco di limoni è stato venduto al mercato</i> 'A bunch of lemons was sold at the marketplace'
1b	Second term	[+ concrete]	<i>Un sacco di limoni sono stati venduti al mercato</i> 'A bunch of lemons were sold at the marketplace'
2a	First term	[- concrete]	* <i>Un sacco di concetti è stato chiarito a lezione</i> 'A bunch of concepts was explained at the class'
2b	Second term	[- concrete]	<i>Un sacco di concetti sono stati chiariti a lezione.</i> 'A bunch of concepts were explained at the class'

Table 1: Possible agreement configurations of Italian quantification expressions with *un sacco di/un mucchio di*

Crucially, while the “semantic” agreement looks always possible, the option of performing the syntactic one seems to be constrained by some semantic properties of the second term of the expression. In fact, basing on the oddness of sentences like (2a) in Table 1, the relevant semantic feature at issue seems to be concreteness. Whereas both agreement configurations are acceptable with concrete nouns like ‘lemons’, only the “semantic” agreement sounds acceptable with abstract nouns like ‘concepts’.

At a first glance, thus, semantic information seems to be effectively interfering with subject-verb agreement. Psycholinguistic models assume that in cases like (1b) and (2b) syntax is overcome by semantics (i.e. Haskell and MacDonald, 2003; Badecker and Kuminiak, 2007; see also Mancini et al., 2014).

Therefore it is desirable to test the acceptability of the different agreement configurations these expressions actually allow with respect to the concreteness of the second term. Such exploration can address the question whether semantics plays a role in the processing of agreement.

Our hypothesis is that only syntax is involved in agreement. Indeed, if syntactic ambiguity should be summoned as the source of *both* agreement options, there would be no need at all to distinguish between a “semantic” and a syntactic agreement.

Following Michaux’s (1992) proposal, we analyse the first term of the expression, i.e. *sacco*, in conditions (1b) and (2b) as a functional element of phrase whose head is the plural noun. However, this analysis is appropriate only for the conditions mentioned, whereas in the remaining conditions the first term of the expression, *sacco*, has a fully

referential meaning; indeed, it cannot be interpreted as a functional element, and it is the head of a phrase that can appropriately work as a probe for agreement.

The present study aims at measuring by means of a self-paced reading test: (i) the acceptability of the four conditions listed in the Table 1, in connection with the syntactic ambiguity hypothesis, and (ii) possible differences between conditions with respect to processing costs, measured at the relevant chunk, that is, the third (i.e. where the inflected verb is shown).

With reference to our first aim, we expect that condition (2a) will receive low acceptability ratings, since syntactic verb agreement with the word *sacco* requires a referential interpretation of the word itself, therefore triggering a semantic mismatch with the following noun when this is [-concrete], as in *Un sacco di sogni è stato tirato fuori dal cassetto* ‘a bag of dreams was taken out of the drawer’.

With reference to our second aim, we expect no differences between conditions to emerge, since we claim there is no competition among different agreement sources (semantic vs syntax), but only one and the same agreement source at work, that is, syntax. In our opinion, this will provide a possible account of the semantic/syntactic agreement issue.

3. *Materials and methods*

40 native Italian speakers, aged 19 to 25 and with 13 to 16 years of education, took part in the study. The experimental trials consisted in 80 passive sentences, 20 per each condition (10 with the quantifier *un sacco di* and 10 with *un mucchio di*). The structure of all the sentences was Noun Phrase—Verb Phrase—Adjuncts (e.g. *Un sacco di limoni—è stato venduto—al mercato*). The second term of each NP was always masculine and plural. The four conditions (see Table 1 above) resulted by the manipulation two variables: the concreteness of the quantified noun (half of them were concrete, the other half abstract), and the probe on which the agreement takes place (the noun or the quantifier, so that half of the verbs were singular, and the rest of them were plural). All quantification expressions and all inflected verbs were balanced with respect to length and frequency by means of it-WaC corpus (Baroni et al., 2009).

140 fillers were also added: condition (f1) presented 40 passive sentences with *un sacco di/un mucchio di*, but in this case the order of the constituents was reversed with respect to the experimental conditions (i.e. VP—NP—Adj., as in *È stato venduto—un sacco di limoni—al mercato*); condition (f2) presented 50 passive sentences like *Molti/Tanti limoni sono stati venduti al mercato* ‘many lemons were sold at the marketplace’, i.e. NP—VP—Adj); finally, condition (f3) presented 50 passive sentences like *Sono stati venduti molti/tanti limoni al mercato* (structure: VP—NP—Adj). Filler sentences were controlled for length, frequency, gender and number of the second terms used (only masculine plural terms) and for the two variables of interest (so half of them displayed plural agreement and the other half singular agreement; half of them had a concrete NP, and the other half an abstract NP). Condition (f2) and (f3) were included to mask the experimental sentences,

Wouldn't You Agree With A Bunch Of Lemons? A Self-Paced Reading Test On
Agreement In Quantification Expressions

whereas condition (f1) was included to prevent the reader from relying on the position of the NP as a cue to solve the syntactic ambiguity.

Both in the case of the experimental items and the fillers, each chunk was formed by a phrase. Such presentation was thought to be more natural in a reading task, especially when compared to a word-by-word presentation. With respect to the experimental sentences, the first chunk was represented by the expressions *un sacco/un mucchio*, the second chunk by a preposition and a masculine noun in the plural form (e.g. *di limoni*), the third chunk by a verb (e.g. *è stato venduto*) and the fourth chunk by an adjunct (e.g. *al mercato*).

Chunks appeared on the screen one at a time. Participants were instructed to press a key to read the following chunk. After having read the last chunk, they had to judge the grammaticality of the whole sentence on a Likert scale ranging from 1 to 4 (1= absolutely not acceptable; 4= completely acceptable).

As stated before, we expect that condition (2a) will receive low acceptability ratings and that no differences in reading times emerge between conditions, since only syntax drives agreement.

4. Results

Data analysis focused on the following dependent variables:

- a) Reading Times at Chunk 3 (the one containing the inflected verb), which was thought to be the critical one, that is, the one responsible of triggering the search for the probe of agreement in the reader's memory;
- b) Rating Scores;
- c) Rating times.

The effect of several predictors on these dependent variables was investigated by means of three mixed effect models, as shown in Table 2, below. The independent variable included in the analyses were length of chunk, frequency of chunk and overall reading time. Variables included in the analyses as random effects were Subject and Item.

Model	Dependent variables	Independent variables
(1)	Reading Time at Chunk 3	Concreteness of NPs; Length of chunk 2; Frequency of chunk 3; Reading Time at chunk 2; Rating Score.
(2)	Rating Score	Concreteness of NPs; Length of the sentence, Reading Time at chunk 4 (Overall Reading Time); Frequency of chunk three.
(3)	Rating Time	Concreteness of NPs; Length of the sentence, Reading time at chunk 4 (Overall Reading Time); Frequency of chunk three.

Table 2: Dependent and Independent Variables in each Model

We analysed the dependent variable of Reading Time at Chunk 3 in Model (1). The analysis showed that the only independent variable influencing reading time at the third chunk was the length of the chunk itself [$\beta = 0.027$, $t = 4.21$]. This is a decisive result, since it implies that experimental condition (Concreteness of the NP) is not linked to reading times.

In Model (2) we analysed the dependent variable of Rating Score. The analysis showed that this variable was related to the Reading Time at chunk 4 (Overall Reading Time) [$\beta = -0.000072$, $t = -4.138$], and to the experimental condition (Concreteness of the NP), with a difference among condition (1b) and condition (1a) [$\beta = 1.20$, $t = 6.975$], condition (2a) and condition (1a) [$\beta = -0.39$, $t = -3.80$], condition (2b) and condition (1a) [$\beta = 1.78$, $t = 7.31$]. These results mean that readers tended to give higher scores to sentences they read more quickly (see also Model 3 below). As shown in Fig. 1, the mean Rating Score for conditions (b) was higher than the one for conditions (a), which means that agreement with the plural term was generally preferred. Moreover, the less acceptable configuration appeared to be condition (2a), as we predicted.

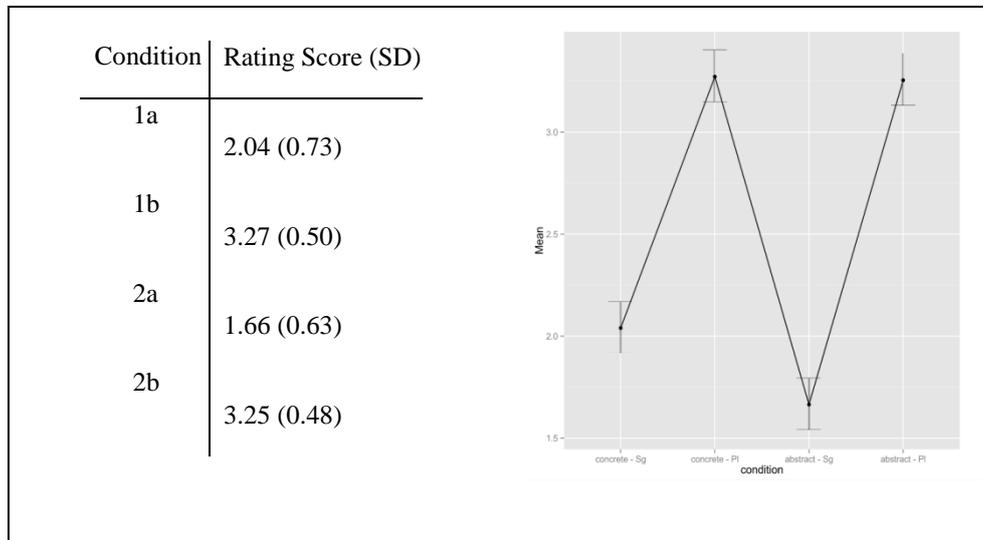


Fig.1: mean Rating Scores and Experimental Conditions

We analysed the variable of Rating Times in Model (3). Independent variables that resulted significant were Overall Reading Time [$\beta = 0.00008$, $t = 4.651$], Rating Score 2 compared to Rating Score 1 [$\beta = 0.49$, $t = 8.45$], Rating Score 3 compared to Rating Score 1 [$\beta = 0.49$, $t = 8.45$], condition (2b) compared to condition (1a) [$\beta = 0.46$, $t = 7.51$]. This data mean that the longer it took to read a sentence, the longer to rate it. Moreover, when the Score was 2 or 3 the Rating Time was longer than when it was 1 or 4, meaning that “absolutely not acceptable” and “completely acceptable” sentences were both read and rated faster than the intermediate conditions. Lastly, it appeared that Rating Time

was shorter in condition (2b) than in condition (1a), which is an unexpected finding that will require further investigation.

To sum up, the analysis through the three mixed effect models showed that the experimental condition (Concreteness of the NP) is not linked to Reading Times, whereas Reading Time at chunk 4 (Overall Reading Time) is linked to Rating Score, meaning that readers tended to give the highest scores to sentences they read most quickly. Also, agreement with the plural term is generally preferred and the less acceptable configuration is condition (2a), as we predicted. Finally, Rating Times and Rating Score are linked to each other, and the more severe the Score (i.e., 1 or 4), the quicker it is given.

5. Discussion

Our first aim was to measure the acceptability of the four experimental conditions. Rating scores revealed a strong preference for the agreement between the verb and the second term of the quantification expression (condition 1b and 2b), whereas lower ratings were obtained by the agreement between the verb and the first term of the quantification expression with an abstract noun (condition 2a), as shown in Fig. 1 above.

These results are consistent with our expectations and with the syntactic ambiguity hypothesis. Indeed, by assuming only a syntactic explanation, it follows that:

- in sentences like (1b) and (2b) *sacco* is actually a functional element of a phrase whose head is the plural noun, and as a functional element it cannot be a probe for agreement. This is the reason why agreement with the second term is preferred and these sentences received high ratings;
- in sentences like (2a) the agreement between the first term, *sacco*, and the verb forces a full referential interpretation of the noun *sacco*, thus driving a semantic implausibility and, as a consequence, a low acceptability rating. Indeed, concepts cannot physically fit in a bag. And this is why this condition was judged ill-formed. Moreover, if this account is true, other semantic features of the noun, not only concreteness, may lead to the same ill-formedness.

Indeed, we would expect a sentence such as *Un sacco di frigoriferi è stato venduto l'altro giorno* 'a bag of fridges was sold the other day' to receive low acceptability ratings, just as well as sentences in condition (2a). This may be also the reason why sentences of condition (1a) were rated higher than those in conditions (2a) but lower than those in conditions (1b) and (2b). Further research is needed on this point.

Results also showed that rating times were shorter when the sentence received polar judgments (i.e. 4 or 1), that is, when it was judged either perfectly acceptable or definitely unacceptable. So, condition (1b) and (2b) on one side and condition (2a) on the other side were judged very quickly. On our account, this result also contribute to identify one and only one agreement source at play: what was definitely good or totally wrong was immediately recognised, with no extra time demanded because of competition among agreement sources. Notably, slower rating times in condition (1a) seem to parallel the fact that sentences in this condition were not judged in a polar way.

Our second aim was to measure possible differences between conditions with respect to processing costs. Results showed a lack of such differences in the reading times at the

critical chunk (the third) in the four experimental conditions; this means that all sentence types required the same time and effort by the reader to be parsed.

This finding allows to argue that no semantic information has interfered or required extra cost during the online processing of agreement. In other words, the parsing of agreement is a purely syntactic matter.

In fact, syntax drives the interpretability of the semantic features of the head involved as a probe for agreement, rather than the other way round. Semantics does not drive the choice of one candidate over another one, nor does it interfere with the process of agreement in some cases and not in others.

Further research on agreement with this method should include a word-by-word presentation paradigm, in order to detect a possible task-dependent bias produced by a chunk presentation, and a task built on sentences like the ones in condition (1a), in order to better investigate reason of the variability in rating scores obtained in this study.

6. Conclusions

The self-paced reading test elaborated in this study explored the way Italian quantification expressions with *un sacco di/un mucchio di* work. These ambiguous phrases seem to permit both a syntactic and a so called “semantic” agreement. However, the results we obtained are consistent with the hypothesis that agreement takes place only at the syntactic level. The so called “semantic” interference is actually better explained as an effect of syntactic ambiguity. We therefore conclude that agreement is an operation handled in syntax.

References

- Badecker, William, and Frantisek Kuminiak. 2007. “Morphology, agreement and working memory retrieval in sentence production: Evidence from gender and case in Slovak”. *Journal of Memory and Language* 56 (1): 65-85.
- Baroni, Marco, Silvia Bernardini, Adriano Ferraresi and Eros Zanchetta. 2009. “The WaCky Wide Web: A Collection of Very Large Linguistically Processed Web-Crawled Corpora”. *Language Resources and Evaluation* 43 (3).
- Bates, Elizabeth, Sandra McNew, Brian MacWhinney, Antonella Devescovi and Stan Smith. 1982. “Functional constraints on sentence processing: A cross-linguistic study”. *Cognition* 11 (3): 245-299.
- Chomsky, Noam. 2000. “Minimalist Inquiries: The Framework”. In *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik* ed. by Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge, MA: The MIT Press.

Wouldn't You Agree With A Bunch Of Lemons? A Self-Paced Reading Test On
Agreement In Quantification Expressions

- Chomsky, Noam. 2001. "Derivation by Phase". In *Ken Hale: a Life in Language* ed. by Michael Kenstowicz, 1-52. Cambridge, MA: The MIT Press.
- Haskell, Todd R., and Maryellen C. MacDonald. 2003. "Conflicting cues and competition in subject-verb agreement". *Journal of Memory and Language* 48 (4): 760-778.
- Haskell, Todd R., Robert Thornton and Maryellen C. MacDonald. 2010. "Experience and grammatical agreement: Statistical learning shapes number agreement production". *Cognition* 114 (2): 151-164.
- Kayne, Richard S. 1989. "Notes on English agreement". In *Parameters and Universals*, Oxford University Press, 2000.
- Mancini, Simona, Nicola Molinaro, Doug J. Davidson, Alberto Avilés and Manuel Carreiras. 2014. "Person and the syntax-discourse interface: An eye-tracking study of agreement". *Journal of Memory and Language* 76 (10): 141-157.
- Michaux, Christine. 1992. "The collectives in French: A linguistic investigation". *Linguisticae Investigationes* 16 (1): 99-124.
- Pollock, Jean-Yves. 1989. "Verb movement, universal grammar, and the structure of IP". *Linguistic Inquiry* 20 (3): 365-424.