

# The syntactic domain of number agreement: attraction effects and pronominal classes

MARIA GARRAFFA\*, ALBERTO DI DOMENICO<sup>o</sup>, ROSALIA DI MATTEO<sup>o</sup>

\**Fondazione Marica De Vincenzi O.N.L.U.S.*,  
<sup>o</sup> *Università degli studi “G. d’Annunzio” Chieti e Pescara*

Assuming that agreement processes are the expression of morphosyntactic relations inside a dedicated projection (an Agr phrase), we examine the feature geometry involved in the construction of subject verb agreement relation. A production experiment was carried out in Italian using the attraction paradigm to investigate the various properties of Subject-Verb agreement when different elements occur in an intervening position between the subject and the verb. Attraction is manifested in Italian, and this occurs more frequently with prepositional phrases as intervener than with object clitics. Furthermore, the agreement configuration in the clitic condition in our experiment had a more complex structure, evidenced by the presence of many errors with all plural forms. The agreement domain of the clitic is sensitive to the nature of the pronoun and this can be seen by comparing different Romance languages. Pronouns have different agreement relations at the marking stage according to whether they are weak pronouns or clitics. The results of the study point to a more finely-tuned model of agreement in which the processing of agreement relations is subject to cross-linguistic variation, given the different types of pronoun found in natural languages.

## 1. Introduction

Agreement is a grammatical operation for establishing relations between elements in a sentence. A central issue in psycholinguistic models concerns the nature of this operation and the different levels of analysis involved. Most research on agreement has focused on the morphosyntactic expression of the number feature, which has two different aspects. One type of number information is notional, and concerns the number of entities in the representation, as in *l’etichetta delle bottiglie* (literally: ‘the label of the bottles’), which refers to more than one entity (labels) even though the plural is not grammatically expressed. The other is purely grammatical and concerns the relationships between words in a sentence. Efficient grammatical processing depends on these two steps being clearly distinct. Notional agreement may play

a role in providing an extra-grammatical strategy for establishing agreement, but this is probably only marginally related to syntactic agreement. Properties that are not strictly syntactic could be expected to disrupt the implementation of agreement but this is not sufficient to suppose that two mechanisms are simultaneously involved in the same operation.

In the present research we focus on the grammatical aspects of morphological agreement in terms of the different structural relations and the different properties of the elements involved in agreement. We are interested in seeing whether mechanisms of verbal agreement are sensitive to grammatical information.

One approach to exploring the factors involved in the AGREE operation is to manipulate the grammatical relationships between the elements in the syntactic configuration. Grammatical agreement in the sentence domain can be investigated using the so-called *attraction* phenomenon (Bock and Miller, 1991), which occurs when an element intervening between two elements sharing an agreement relation and mismatched in the relevant features attracts agreement. This is simplified in the configuration in (1).

- (1)     [...X, Singular] [...Z, Plural] [...Y, Singular]  
           \*The boat of the American soldiers leave in the fog

There is interesting evidence that there is an asymmetry between singular and plural: plural local nouns co-occurring with singular head nouns have been found to give rise to greater interference effects both in production errors (Bock and Miller, *ibid*) and in slower reading times in comprehension (Pearlmutter et al. 1999). These studies consistently show between 4% and 11% agreement errors in production in attraction configurations.

The phenomenon of “attraction” seems then to be asymmetric, which can be accounted for within a formalism that assumes syntactic features to be binary, possessing either a marked value of a given property, in this case (+ number), or an unmarked value (- number) (Jakobson, 1957). According to the “Marking and Morphing” model (Bock and Eberhard, 1993), the asymmetry arises at the morphological level, which is responsible for morphemic building of speech output, and is due to interference within a “reconciliation” process linking the marking stage, assumed to be notional in nature, and the lower-level Morphing stage<sup>57</sup>.

A similar explanation for comprehension data has been put forward by Pearlmutter et al. (1999), who assumes that comprehension difficulties in sentences with singular head and plural local nouns arise from inadvertent head-overwriting on an on-the-fly NP number computation rather than from a backtracking mechanism, leading to speculation that number is syntactically considered only when a plural feature is encountered.

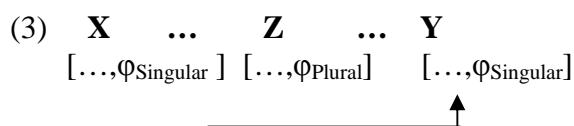
<sup>57</sup> The fact that asymmetry between singular and plural has always been tested in this particular environment, where structural locality and linear precedence conflict, leaves open many possible explanations for the level of processing at which the feature asymmetry originates, in particular whether the need for notional assignment of number to phrases plays a role.

*The syntactic domain of number agreement*

Interesting data have come from studies on the impact of syntactic factors on agreement errors using the attraction paradigm in English. Bock and Cutting (1992) observed that there are significantly fewer attraction effects in 2b than in 2a.

- (2) a. [The editor of the history books] is/are...  
 b. [The editor] who rejected the books is/are...

This experiment suggests that there is considerable processing difficulty in regulating the relations between elements which are structurally similar and are part of the same syntactic phrase, as in 2a, confirming the view that constituents are the relevant unit over which agreement takes place. In 2b, the intervening phrase, *the books*, is part of the relative clause. To understand all this, let's suppose that the processor, having to proceed under time pressure and with strong working memory constraints, wants clear indication of the nature of the elements to link in a local relation. When linear and hierarchical orders conflict, the system is prone to error. These results can be interpreted as a consequence of minimality effects induced by the agreement feature of the intervening prepositional modifier, as schematized in (3).



It is reasonable to suppose that in a situation where the processing cost is high (maintaining two potential agreement relations before the verb form) the system is prone to error: A marked feature activates the agreement operation and attraction arises in a minimality environment.

There is evidence attesting to subject modifiers occasionally interfering with verb agreement when there is a number mismatch with the head noun. However, number features within the VP have rarely been studied. A recent study by Franck et al. (2007) on the structural properties of agreement revealed different effects with different syntactic relations. Attraction effects were investigated in French with two different classes of interveners: prepositional modifiers (4a) and clitic object pronouns (4b).

- (4a) [ Subj... N<sub>head</sub>...N<sub>mod</sub>... ] AgrS  
 Le professeur-SG des élèves-PL lis-SG/\*lisent-PL  
*The professor of the student reads/\*read*
- (4b) [ Subj...N<sub>head</sub>... ] [AgrOP N<sub>clitic</sub>] AgrS  
 Le professeur-SG les-PL lit-SG /\*lisent -PL  
*The professor them reads/\*read*

The main finding of this experiment was that there was a significantly greater number of attraction errors with clitics, as in 4b, than with prepositional modifiers, as in 4a. They put this down to differences in the structural relations involved, given that object clitics intervene not only linearly but also hierarchically as they are in an argument position. The study with French adults seems to indicate that the mechanism for implementing agreement is influenced

by the nature of the local constituents, with more attraction effects occurring with an object clitic intervener. It was also found that in the clitic condition plural head nouns generated more errors than singular head nouns, and plural intervening elements more than singular interveners.

Other interesting data come from Spanish. Crucial for the present investigation, manipulation of clitic object pronouns does not give rise to attraction effects in Spanish (Anton-Mendez, 1996), where the local elements were object clitics marked for accusative case, a null-effect was found. Spanish differs from other Romance languages in that it lacks participial agreement and therefore the head for object agreement. The Spanish data indicate that differences between languages may be related to grammatical selection and that a non-active agreement position in a language could not induce attraction effects. A possible explanation for this is that the clitic itself is not relevant for attraction effects. In French, as in Italian but not Spanish, the clitic pronoun is an active element for agreement. This can be seen in past participle agreement with a preverbal object clitic.

- (5a) Les pommes, je **les** ai mangées  
 (5b) Le mele **le** ho mangiate  
 (5c) Las manzanas **las** he comido [No Agr]  
*The apples, I them have eaten*

The absence of attraction effects with an object clitic in Spanish can be ascribed to the absence of a fully-fledged AgrOP intervening between the subject and the verb. The assumption behind this argument is that the object agreement phrase is the potential antecedent causing attraction effects, not the clitic element per se or the particular relations involved in cliticization.

### 1.1 The grammatical properties of subject-verb agreement

The agreement case studied here is Subject-Verb predication, which is governed by morphological variation related to syntactic and/or semantics factors. Subject-verb agreement for number can be modified according to:

i) the nature of the subject: phrasal subjects, as in (6), as opposed to non-phrasal subjects, as in (7). This is exemplified in the different implementations of agreement found with coordination (Heycock and Zamparelli, 2005; Piccolo, 2002):

- (6) Che sia partito tardi e che sia tornato presto è/\*sono un fatto molto strano  
*That he left late and came back early is/\*are a strange thing*  
 (7) Carla e Maria sono/\*è buone compagne di viaggio  
*Carla and Maria are/\*is good travelling companions*

ii) the nature of the predicate: distributive predicates, as in 9, as opposed to non-distributive predicates:

- (8) I due ragazzi cantano e ballano  
*The two boys sing and dance*

*The syntactic domain of number agreement*

(9) Carla e Maria sono un clown e un saltimbanco  
*Carla and Maria are a clown and an acrobat*

iii) subject position and its relationship to agreement, exemplified in the difference between preverbal (10) and postverbal subject agreement (11), or in terms of structural distance as in (12) and (13) :

(10) Trois filles sont arrivées

(11) Il est arrivé trois filles

(12) The queen of England is you

(13) The queen of England may be you (from: Heycock and Kroch 1999)

The present research deals with the third point. We explore some of the properties related to structural distance and examine different intervening elements. Following Franck et al. (see (4)), we investigated the phenomenon of attraction by comparing linear intervention inside the same phrase, as in the classical attraction test with prepositional modifiers (14), with intervening elements such as object clitics (15), also linear interveners in subject-verb agreement.

(14) [ Subj... N<sub>head</sub>...N<sub>mod</sub>... ] AgrS  
Il professore-SG degli studenti-PL legge-SG/\* leggono-PL  
*The professor of the student reads/\*read*

(15) [ Subj...N<sub>head</sub>... ] [AgrOP PRO<sub>clitic</sub>] AgrS  
Il professore-SG lePL legge-SG /\*leggono-PL  
*The professor them reads/\*read*

The aim of the study is first of all to explore the attraction paradigm in Italian and to see whether we can make finer distinctions in the grammatical elements involved. We are also interested in obtaining real-time processing data from different conditions in order to assess potential complexity effects involved, for example, in the object clitic condition.

Given that agreement is expressed differently across languages according to their grammatical properties, we stress the importance of cross-linguistics studies for a better understanding of the role of grammatical elements in the implementation of agreement. In particular, we expect a greater number of errors in the condition with object clitics, as in French, if c-command is the agreement domain. New data have the potential to open new discussions on the structural relations involved in attraction.

### 3. Methods

#### 3.1 Participants

One hundred and five undergraduate students (88 female) participated in the experiment. Their mean age was 22.1 years, and their age range 20–35 years (standard deviation 2.03). They were all native Italian speakers with normal or corrected-to-normal vision.



*The syntactic domain of number agreement*

items (6 of each kind, with a correction of two) and 50 filler items. 25 of which contained a prepositional modifier with an indeclinable noun, as in (18), and 25 contained a reflexive pronoun, as in (19).

- |      |   |                                   |
|------|---|-----------------------------------|
| (18) | Il maglione di lana<br><i>The sweater of wool</i> | INGRASSARE<br><i>TO GET FAT</i>   |
| (19) | Il dottore si<br><i>The doctor himself</i>        | AMMALARE<br><i>TO BECOME SICK</i> |

### 3.3 Procedure

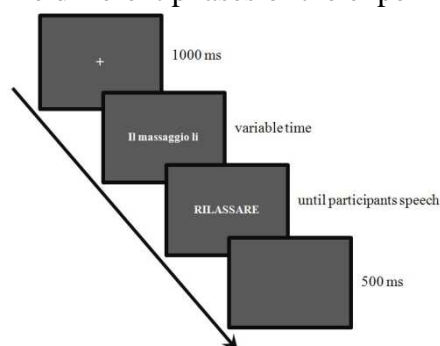
All participants were tested individually. They were presented with a short list of practice items before beginning the experiment in order to familiarize them with the task.

A fixation point appeared in the center of the screen for 1000 ms followed by the first part of each sentence for an interval of time varying with the length of the sentence in order to ensure consistent processing time (sentences with a prepositional modifier were always shorter than sentences with a clitic object pronoun). The interval was calculated with the variable Serial Visual Presentation (vSVP) formula [(187 ms x sentence word number) + (27 ms x sentence character number, including spaces)] developed by Otten & Van Berkum (2008). Then the verb appeared and remained on the screen until the subject had verbally produced a singular or plural finite form of it. During the practice session participants were trained to pronounce each verb clearly, without hesitation, and, most importantly, without drawling the last part of the word (in Italian the verb ending provides number information). This procedure was adopted to ensure that sentence recognition was completed before participants started to say the verb, thus ensuring that the two tasks were not carried out simultaneously and that differences in speech onset times really reflected differences in processing cost. Once the verb form was spoken, an empty screen appeared for 500 ms before the fixation point reappeared (the procedure is schematized in Figure 1).

Sentences were presented randomly and each experimental session (Practice + Experimental Session) lasted approximately twenty minutes.

**Figure 1**

The different phases of the experiment are reported along the time-line.



Verbs production times were measured using a special microphone connected to E-Prime software. During each experimental session verbal responses were

registered with a second microphone connected to a second PC. This allowed us to verify verb form accuracy and the presence of hesitations.

### 3.4 Data analysis

Responses with hesitations or drawled endings were not included in the analyses. In addition, a threshold of three standard deviations was used to filter out production time outliers, and mean production times were computed only for correct singular or plural responses.

## 4. Results

A total of 205 (8.15%) errors in the prepositional modifier condition and 227 (8.98%) in the clitic object pronoun condition were produced. Error distribution is shown in Table 2.

**Table 2**

Raw data and % errors in the various experimental conditions

CONDITION	FORM			
	SS	SP	PP	PS
PREP	26 (4,07%)	110 (17,55%)	23 (3,53%)	46 (7,45%)
CLIT	17 (2,58%)	68 (10,95%)	81 (12,75%)	61 (9,64%)

### 4.1 Production times

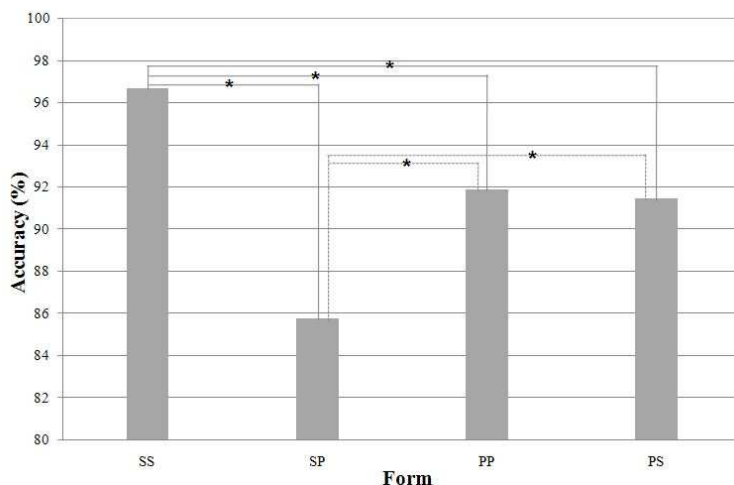
An ANOVA with production time as the dependent variable and the factors *Condition* with two levels (PREP vs. CLIT) and *Form* with four levels (SS, SP, PP and PS) did not show any significant effects.

### 4.2 Production accuracy

An ANOVA with production accuracy as the dependent variable and the factors *Condition* with two levels (PREP vs. CLIT) and *Form* with four levels (SS, SP, PP and PS) showed a significant effect of *Form* ( $F_{3,312} = 24.75$ ,  $p < 0.001$ ). This reveals that participants were generally more accurate in producing the correct verb in SS sentences than in SP, PP and PS sentences (post-hoc Tukey test  $p < 0.001$ ), while their performance was less accurate in SP sentences than in the other sentence forms (post-hoc Tukey test  $p < 0.001$ ) (see Figure 2).

**Figure 2**

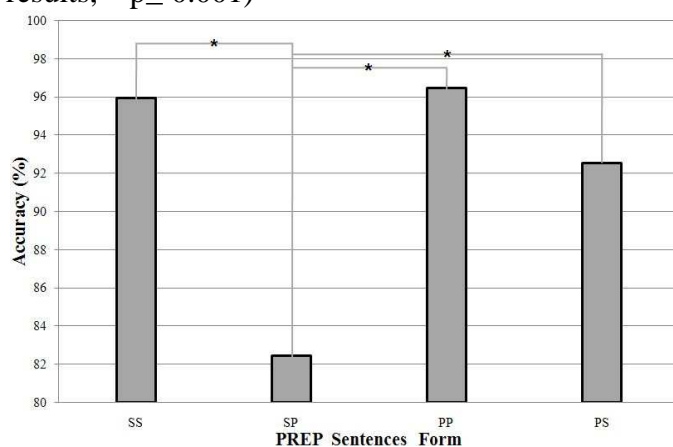
Mean % accuracy with the various sentence forms (asterisks indicate significant post-hoc test results, \*  $p \leq 0.001$ )



The *Condition x Form* interaction was also significant ( $F_{3,312} = 18.86$ ,  $p < 0.001$ ). To understand this effect better, we conducted a separate ANOVA for each Condition with production accuracy as the dependent variable and the factor *Form* with four levels (SS, SP, PP and PS). A significant effect of *FORM* in the PREP Condition ( $F_{3,312} = 30.39$ ,  $p < 0.001$ ) showed that participants were less accurate in producing verbs in the PREP SP condition than in the other PREP conditions (post-hoc Tukey test  $p < 0.001$ ) (see Figure 3), whereas a significant effect of *Form* in the CLIT Condition ( $F_{3,312} = 14.18$ ,  $p < 0.001$ ) showed that participants were more accurate in producing verbs in CLIT SS sentences than in the other CLIT sentences (post-hoc Tukey test  $p < 0.001$ ) (see Figure 3).

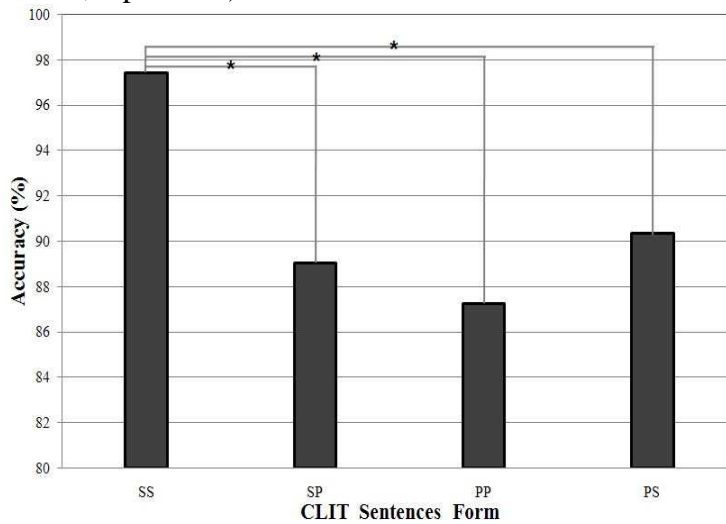
**Figure 3**

Mean % accuracy with PREP forms (asterisks indicate significant post-hoc test results, \*  $p \leq 0.001$ )



**Figure 4**

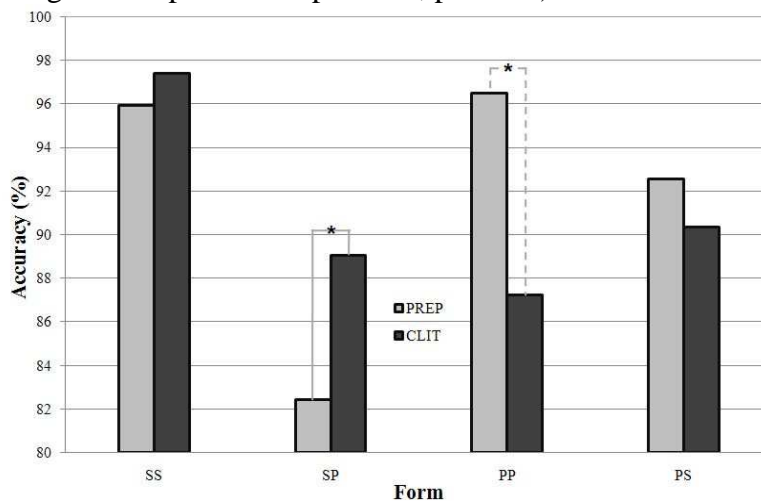
Mean % accuracy with CLIT forms (asterisks indicate significant post-hoc test results, \*  $p \leq 0.001$ )



Finally, paired comparisons showed that there were no differences in accuracy between PREP SS and CLIT SS sentences and between PREP PS and CLIT PS sentences, but participants were more accurate in producing verbs in CLIT SP sentences than in PREP SP sentences ( $p < 0.001$ ), and in producing verbs in PREP PP sentences than in CLIT PP sentences ( $p < 0.001$ ) (see Figure 5).

**Figure 5**

Mean % accuracy as a function of form and condition (asterisks indicate results of significant paired comparisons,  $p \leq 0.001$ )



## 5. Discussion

Our experiment using the attraction paradigm in Italian produced some very interesting data. Attraction, calculated as a significant difference between accuracy with the asymmetric form SP and the baseline (the symmetrical unmarked form SS), was measured in both prepositional and clitic conditions. Both SP forms differed from the baseline and this effect can be clearly seen in figures 3, 4 and 5. There were also slight differences between the clitic and

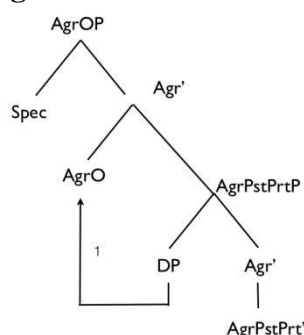
prepositional conditions. Firstly, comparison of these two conditions in the asymmetric form (SP) showed there was greater accuracy, so less attraction, with a clitic intervener than with a prepositional intervener. Secondly, all forms in the clitic condition differed from the baseline (SS) but there were no significant differences between the PP, PS and SP forms. This pattern was not found with prepositional modifiers, showing their status to be different to that of clitics. In the following paragraph we discuss the two main aspects of the results.

### 5.1 The status of object clitics in Italian

Our initial hypothesis was that in Italian attraction effects in subject-verb agreement when the intervener is an object clitic should be more similar to the effects found in French than those in Spanish. This is because French and Italian have object agreement but Spanish doesn't (see the examples in (5)). Instead, object clitics in our experiment with the attraction paradigm revealed a kind of intermediate pattern for Italian. More precisely, unlike Spanish, attraction occurs when a plural object clitic intervenes between a subject and the verb. However, comparison of the two conditions shows that in Italian, unlike in French, there is a less pronounced effect with object clitics than with prepositional modifiers. To understand this intermediate pattern in Italian better, it is necessary to adduce a new element into the discussion on grammatical agreement. An explanation needs to be found for the fact that we found fewer attraction effects with object clitics than with prepositional modifiers in Italian while Franck et al. found the opposite pattern in French. We take the view that cross-linguistic differences play an important role in agreement operations. The attraction paradigm is a useful instrument for testing subtle distinctions between languages in grammatical processing.

Romance languages have rich pronominal systems, which have been extensively investigated in linguistic theory (for an overview see Cardinalletti and Starke, 1999). Pronouns divide into three different classes according to their grammatical properties: clitics, weak pronouns and strong pronouns. These categories are fairly fluid and the number of pronouns in each class differs according to languages. Simplifyng, clitics are nominal arguments closely linked with the verbal domain. In Italian, the pronominal argument moves to AgrO, the landing site for clitics, as a head, as illustrated in fig.2 (see Belletti, 1999 for all formal details).

**Fig. 2**



Belletti (ibidem) argues that the status of object clitics is different in French. In





implemented similarly in Italian and French in the condition with a clitic intervener. We now have to reconsider this hypothesis and add a new level to the analysis of this pronominal category.

As table 3 shows, object pronouns have a different status in French and in Italian, which accounts for greater attraction effects in French.

**Tab. 3**

	<b>AgrPstPrt</b>	<b>Weak Pronoun</b>	<b>Object clitic interference</b>
ITALIAN	+	-	-
FRENCH	+	+	+
SPANISH	-	-	-

In Italian, on the other hand, there is no proximity between the clitic, an X°, and the subject, an XP, as illustrated in 27.

- (27)
- |               |     |               |     |             |  |
|---------------|-----|---------------|-----|-------------|--|
| X             | ... | Z             | ... | Y           |  |
| [...,-num]]XP |     | [...,+num]]X° |     | [...,-num]] |  |
| L'arbitro     |     | li            |     | fischia     |  |

### 5.2 Classification of pronouns

Like French, clitics in Italian appear to have a stronger disruptive effect in subject-verb agreement. Italian does not seem to display asymmetrical interference and when either or both subject and object clitics are plural in Italian there is the same low degree of interference. Clitics have a different mechanism related to their specific status. Many errors have also been found with PP forms in French (Franck et al., 2006) and in German (Hemforth et al., 2003). It is not clear, given that both elements bear a marked + number feature, why the verb can appear in a singular form. Fortunately, qualitative analysis of the errors throws some light on this apparently unclear situation. In four cases we found with clitics in PP form a person mistake, so an insertion of a new subject as in (28).

- (28) I giocatori li canto *The player them I sing*

- (29)
- |                           |     |                           |     |                    |      |
|---------------------------|-----|---------------------------|-----|--------------------|------|
| X                         | ... | Z                         | ... | <pro>              | Y    |
| [...,+gen,+num,3°pers]]XP |     | [...,+gen,+num,3°pers] X° |     | [...,-num,1°pers]] |      |
| The players               |     | them                      |     | I                  | sing |

In a generic experimental task without a preamble, as the one described in methods section, the clitic must search for a suitable referent to be integrated in the sentence. If subject and clitic have the same features (number and gender) this is a clear indication to reanalyze the first noun as a topic object rather than a subject and to insert *pro* in order to establish the necessary subject-verb relation, as in 29.

Different processing models make different predictions about the effect of distance on integration. Revision of the entire fragment does not seem to be economical according to standard processing models, such as Gibson's (1998). In Gibson's model, each incoming word activates the information associated with it in the mental lexicon. The level of activation may decrease as subsequent

#### *The syntactic domain of number agreement*

material is processed, and the information is harder to maintain when subsequent material introduces new discourse referents (Gibson and Warren, 1997). Therefore, when additional discourse referents are introduced between a (non-matrix) subject and its verb, it is harder to integrate the verb with the subject, because more effort is needed to reactivate the information associated with the subject. This model therefore predicts increasing processing difficulty with the verb as distance (in terms of the number of new discourse referents) between the subject and the verb becomes greater, even when the subject features have been correctly tracked.

This prediction is falsified by our results. In particular, a matching of features between a noun and the following pronoun can be a stronger interpretation source. Clitics need a clear antecedent and the first noun is the easiest candidate. At this point the system prefers to insert a new referent and revise the “marking stage” so that the first noun is reinterpreted as a topicalized object. A stronger principle of Grammatical Codification of these structures is at play to consider these as stable structures and to be established quickly grammatical relations.

## **6. Conclusions**

The attraction effect is the result of an analysis of syntactic structure applied to abstract representations. We found attraction to be sensitive to categorial distinctions and to indicate subtle cross-linguistic distinctions in the pronominal systems of Romance languages. Fewer attraction effects were found with clitics than with prepositional modifiers, since there are no weak subject pronouns in Italian and clitics must be in a local configuration with the verb. According to minimality effects these elements are different enough to be distinguished from subjects.

We also found that a different mechanism operates in the construction of pronominal arguments. The possible interplay between marking and morphing levels should be better investigated to develop synergies between what it is called grammatical agreement (morphing level) and discourse level information (marking level). As for the present research with Italian object clitic the construction of grammatical relations should be investigated in additional languages in order to increase our understanding of the agreement operation.

## **References**

- Anton Mendez, I., 1996 *Clitics and attraction errors* M.A. thesis University of Arizona
- Bock, J. K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology*, 23, 45-93.
- Bock, K. and Eberhard, K.M. ,1993 The meaning, sound, and syntax of English number agreement. *Language and Cognitive Processes*, 8, 57-99.
- Bock, K., and Cutting, J. C., 1992 Regulating mental energy: Performance units in language production. *Journal of Memory and Language*, 31, 99-127
- Frank, J., U.H. Frauenfelder and Rizzi, L., 2006 *Agreement and movement: syntactic analysis of attraction* Cognition: vol.101,pp.173-216.
- Cardinaletti, A. and Starke, M. The typology of structural deficiency on the three grammatical classes. A case study of the three classes of pronouns. In: van Riemsdijk, H. (Ed.), *Clitics in the Languages of Europe*. de Gruyter, Berlin, pp. 145–233.
- Gibson, E. 1998 Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68, 1-76.
- Kayne, R., 1991 Romance clitics, verb movement and PRO. *Linguistic Inquiry* 22.4: 647-86.

- Hamann Cornelia / Belletti Adriana, 2008, Developmental patterns in the acquisition of complement clitic pronouns: comparing different acquisition modes with an emphasis on French, *Rivista di Grammatica* 31: 39-78.
- Heycock, C. and A. Kroch (1999). Pseudocleft connectedness: Iinterface. *Linguistic Inquiry* 30, pp. 365-397.
- Heycock, C. and Zamparelli, R., 2005, Friends and colleagues: Plurality, coordination, and the structure of DP, *Natural Language Semantics* 13: 201-270.
- Pearlmutter et al., 1999 Linear versus hierarchical agreement feature processing in comprehension. 12<sup>th</sup> CUNY conference, NJ.
- Otten, M., and Van Berkum, J. J. A. (2008). Discourse-based lexical anticipation: prediction or priming? *Discourse Processes*, 45(6), 464-496.