

Prominence in coreference processing: weight of syntactic function and order of mention

Paula Luegi¹, Marcus Maia², Armanda Costa¹

¹University of Lisbon, ²Federal University of Rio de Janeiro

paulaluegi@fl.ul.pt, maiamarcus@gmail.com, armandacosta@fl.ul.pt,

AMLaP

Paris . France . September 1st-3rd . 2011

1. Introduction

Common ground about anaphoric processing:

- The choice of an antecedent for a pronoun is based on the prominence of the available entities in the discourse: the more salient an antecedent is in the discourse the less marked and informative (more reduced) will be the anaphoric expression referring to it (Ariel, 1996)

Matter of debate:

- What defines the prominence of an entity?
 - Some theories propose that only one factor is responsible for prominence ascription
 - Other theories adopt a multifactor perspective

Current stage:

- Many studies have demonstrated the importance of different types of linguistic information such as Syntax (Subjecthood, Order of mention/Topicality, Focus Information), Semantics (Implicit causality, Agenthood) or Pragmatics
- However few studies have combined and contrasted those types of information against one another. For instance, many authors defend that:
 - Syntactic function or Subjecthood is the most important factor for saliency ascription
 - Anaphoric expressions like Null Pronouns are preferred to retrieve a sentential SUBJECT as antecedent while Overt Pronouns are preferred to retrieve NON-SUBJECT constituents (Costa, Faria, & Matos, 1998, and Costa, Faria, & Kail, 2004, for European Portuguese; Corrêa, 1998, and Melo, & Maia, 2005, for Brazilian Portuguese; Carminatti, 2002, for Italian; Alonso-Ovalle, Clifton, Frazier, & Solera, 2002, for Spanish)
 - Others claim that Order of Mention is the factor that defines the saliency of an antecedent:
 - First mentioned entities are very prominent, regardless of their syntactic status (Gernsbacher, & Hargreaves, 1988, for English; Carreiras, Gernsbacher, & Villas, 1995, for Spanish)
- While in the former studies the SUBJECT was always the first mentioned entity, in the latter studies the alleged preference for first mentioned entities was not tested with pronouns
- Some factors that have been contrasted in different languages:
 - Gender information and Accessibility (Arnold, Eisenband, Brown-Schmidt, & Trueswell, 2000, in English)
 - First-Mention and Subject-Preference (Järvikivi, Gompel, Hyönä, & Bertram, 2005, and Kaiser, & Trueswell, 2008, in Finnish; Kaiser, & Vihman, 2006, in Estonian)
 - Topic and Focus (Kaiser, 2006, in English)

In general, these studies have shown that many factors contribute to saliency ascription. However, some aspects may not have been deeply explored due to the characteristics of the tested languages. For instance, English is neither a free word order language nor a pro-drop one (no null SUBJECTS are allowed) and, in the studies with Finnish and Estonian, free word order languages, only overt pronouns and demonstratives were used to test pronoun resolution preferences.

Note: We use the terms 'prominence' and 'saliency' as synonyms, to refer to entities that are the centre of attention at a point in the discourse.

AIM OF THE PRESENT STUDY

Contrast the weight of syntactic function and order of mention on prominence ascription during pronoun resolution. To do so we tested Portuguese complex sentences with SVO or OVS main clause orderings, followed by a juxtaposed temporal adverbial subordinate clause. The SUBJECT of the subordinate clause is either a null or an overt ambiguous pronoun that must be identified by an antecedent in the previous clause.

2. Experiment

Participants: 24 adult native speakers of European Portuguese from the University of Lisbon

Stimuli: 20 experimental items in 4 conditions ((1)-(4)) + 40 fillers

(1) Main clause SVO order + null pronoun (SVO+NULL)

O mecânico trabalhou com o engenheiro na oficina quando *o* remodelou o carro de competição.
The mechanic_{SUB} worked with the engineer_{OBL} in the garage when [he] remodeled the competition car.

(2) Main clause OVS order + null pronoun (OVS+NULL)

Com o engenheiro trabalhou o mecânico na oficina quando *o* remodelou o carro de competição.
With the engineer_{OBL} worked the mechanic_{SUB} in the garage when [he] remodeled the competition car.

(3) Main clause SVO order + overt pronoun (SVO+OVERT)

O mecânico trabalhou com o engenheiro na oficina quando ele remodelou o carro de competição.
The mechanic_{SUB} worked with the engineer_{OBL} in the garage when he remodeled the competition car.

(4) Main clause OVS order + overt pronoun (OVS+OVERT)

Com o engenheiro trabalhou o mecânico na oficina quando ele remodelou o carro de competição.
With the engineer_{OBL} worked the mechanic_{SUB} in the garage when he remodeled the competition car.

Procedure: Visual World Paradigm: participants listened to sentences while viewing pictures with the representation of the two mentioned characters and with the location of the action described in the main clause (eye movements were registered, at a 500Hz rate, with an IVIEW X™ HI-SPEED) + Questionnaire: final interpretation question

Analysis:

Visual World Paradigm (on-line task): percentage of looks to each entity (SUBJECT, OBJECT or LOCATIVE) starting 200ms after the onset of the pronoun, on the overt pronoun condition, or after the onset of the Verb, on the null pronoun condition, and for a 1400ms time window. Only percentage of looks to SUBJECT and to OBJECT are contrasted

Questionnaire (off-line task): percentage of SUBJECT/OBJECT choice

Note: The term OBJECT is used in a general sense, referring to Obliques prepositional phrases. In our conditions, when the PP is moved to the beginning of the sentence, the preposition will work as a cue to the parser to point out that it is not the Subject that is in the first position, preceding the verb, but a complement or an adjunct.

HYPOTHESES & PREDICTIONS

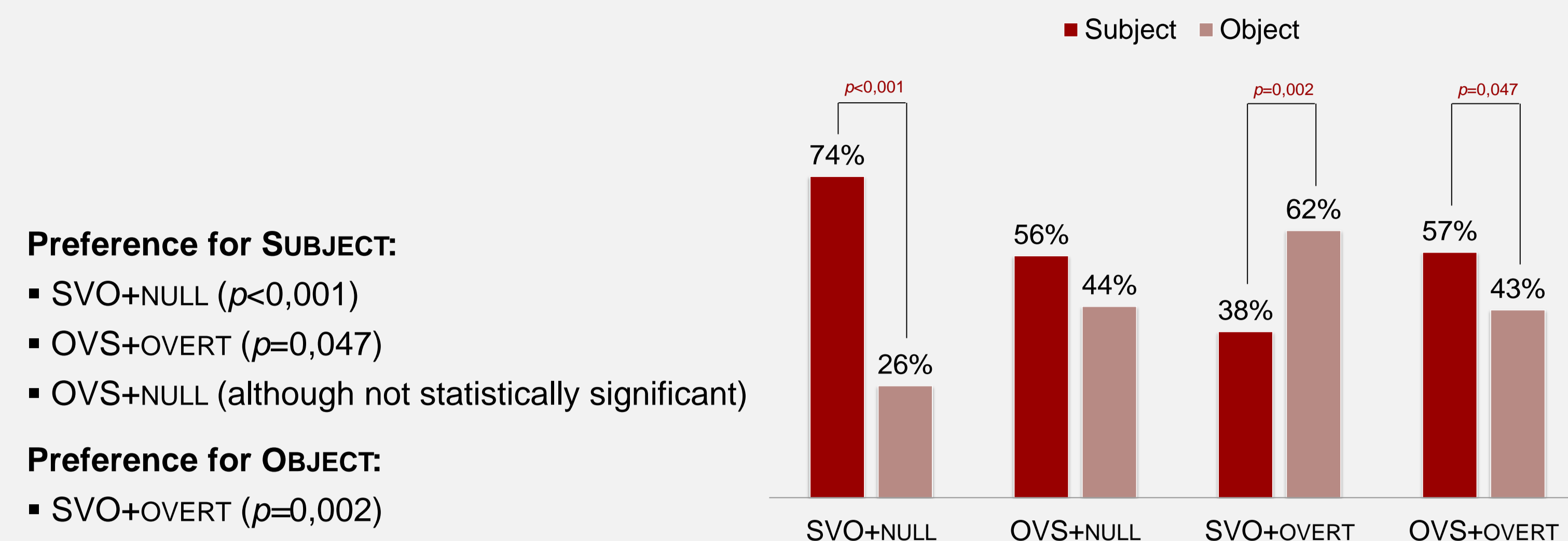
H1: Prominence depends only on Syntactic-function information, hence the null pronoun will retrieve always the Subject and the overt pronoun will be preferred to retrieve the Object.

H2: Prominence depends on the Order of Mention information, and so the null pronoun will prefer the entity that is first mentioned and the overt pronoun will prefer the second mentioned entity.

H3: Prominence is the result of the combination of various factors, namely Syntactic-function and Order of Mention. Therefore there will be different results depending on the combination of different linguistic information, for instance, null pronoun-Subject relationship will not be as straightforward as expect if Subject is not the first-mentioned entity, or overt-Object, if the Object is the first-mentioned entity.

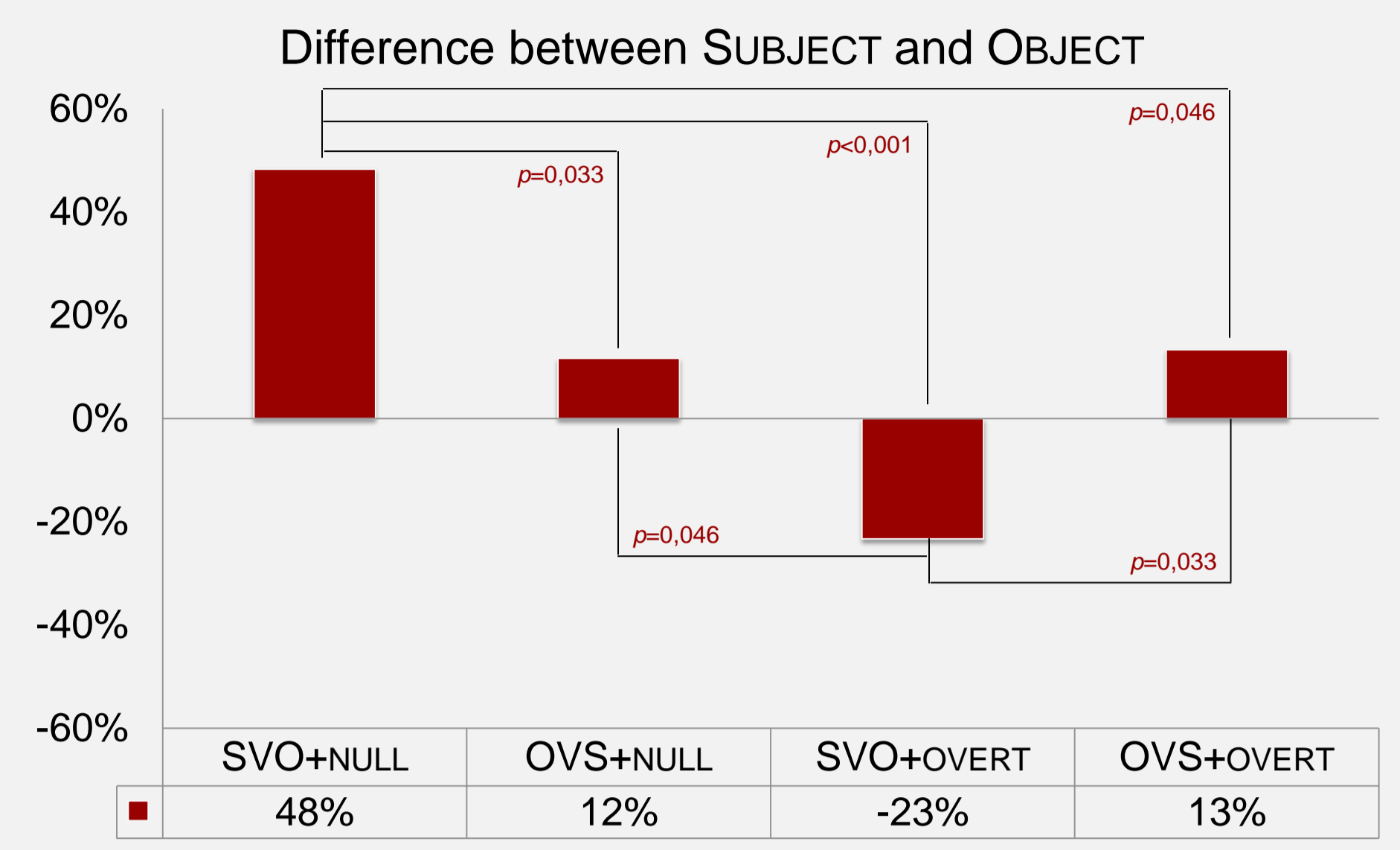
3. Results

3.1. Questionnaire

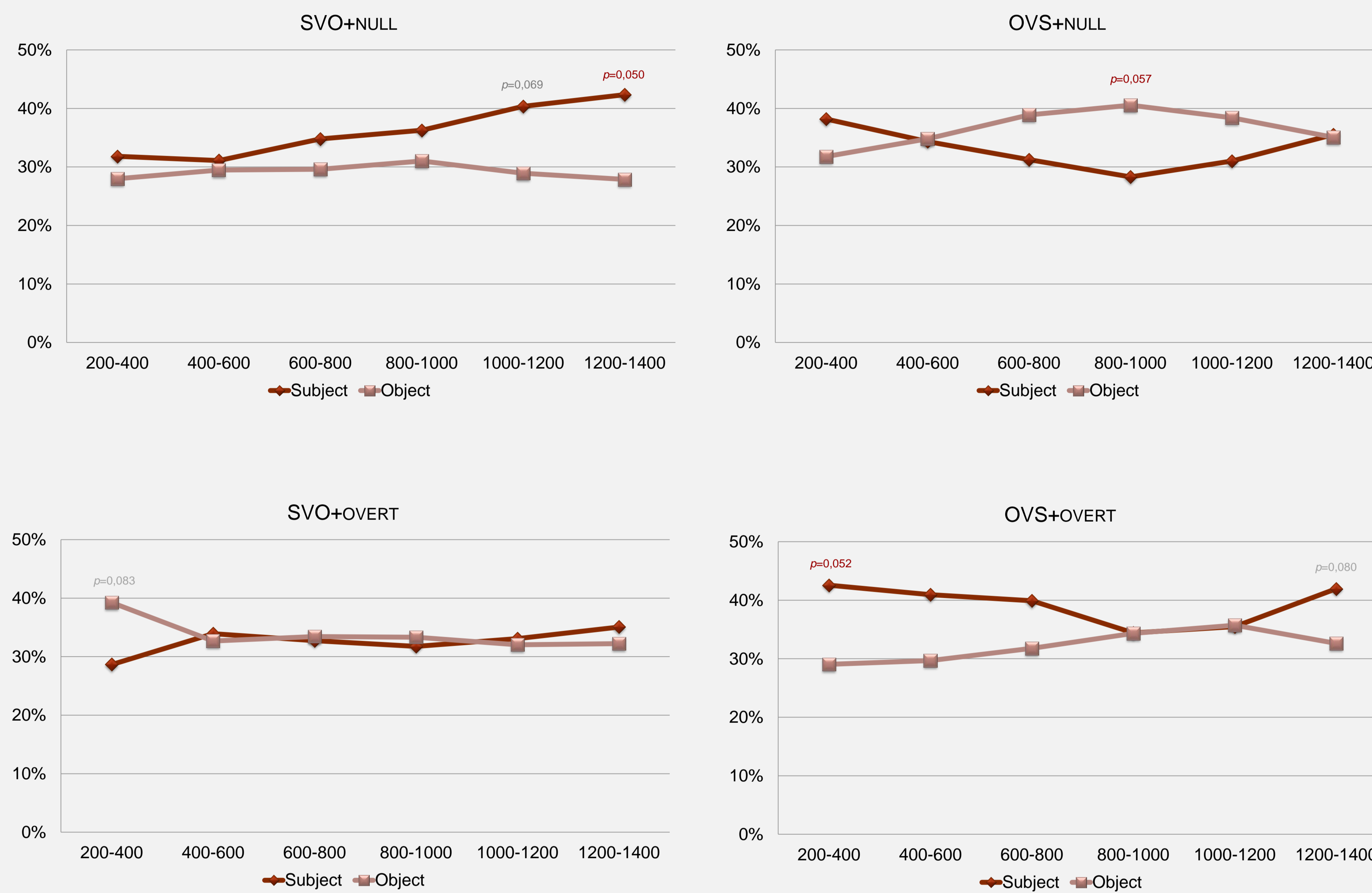


Differences between conditions (SUBJECT choice minus OBJECT choice)

- SVO+NULL and SVO+OVERT are different from OVS+NULL and from OVS+OVERT and also different from each other



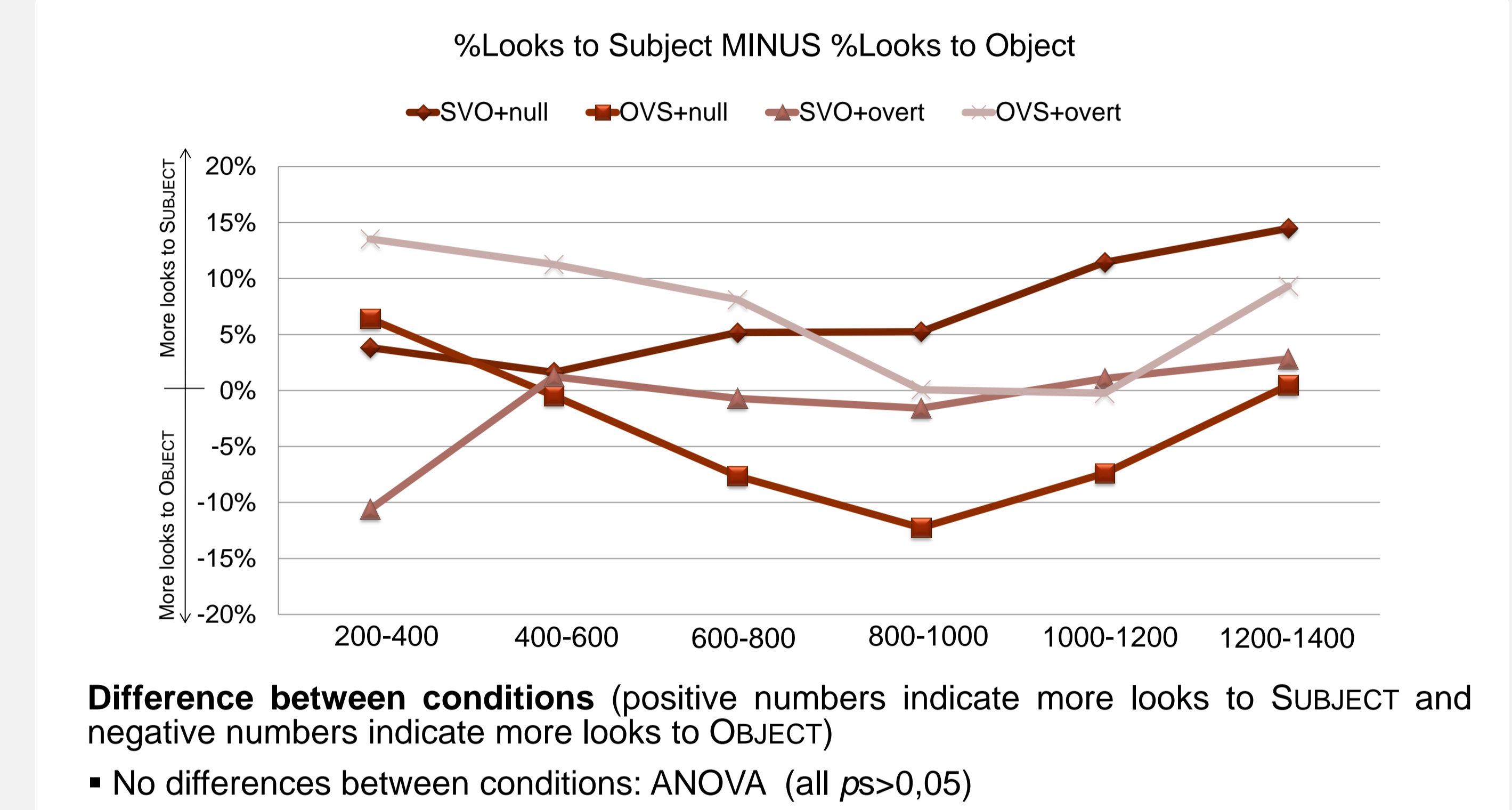
3.2. Visual World Paradigm



Percentage of looks: SUBJECT versus OBJECT

- SVO+NULL: more looks to SUBJECT, greater at 1000ms-1200ms ($*p = 0.069$) and at 1200ms-1400ms ($p = 0.050$)
- OVS+NULL: more looks to OBJECT, greater at 800ms-1000ms ($p = 0.057$)
- SVO+OVERT: slightly initial preference for object but not statistically different
- OVS+OVERT: more looks to SUBJECT, greater at 200ms-400ms ($p = 0.052$)

In sum, more looks to SUBJECT in SVO+NULL and OVS+OVERT conditions and more looks to OBJECT in OVS+NULL condition. No preferences in SVO+OVERT condition.



4. Discussion

Results are summarized in the following table:

	Questionnaire	Visual World Paradigm
SVO+NULL	SUBJECT	SUBJECT
OVS+NULL	SUBJECT _{noSig}	OBJECT
SVO+OVERT	OBJECT	OBJECT _{noSig}
OVS+OVERT	SUBJECT	SUBJECT

- Only in the SVO+NULL condition there is a clear preference to retrieve the SUBJECT with the null pronoun. When the SUBJECT is also the first-mentioned entity there is a straightforward preference to retrieve it with the null pronoun
- In the OVS+NULL condition there is a preference to choose the SUBJECT as the antecedent of the null pronoun in the off-line task (although not statistically significant) but the preference goes for the OBJECT during the on-line task. This result seems to indicate that the order of mention of the SUBJECT influences its saliency and that an OBJECT that is in a topic position may become more salient than a post-verbal SUBJECT, making it a good antecedent for the null form of the pronoun
- The former hypothesis may also explain the results of OVS+OVERT condition (preference for SUBJECT). That is, if the SUBJECT in post-verbal position is not as salient as an OBJECT in topic position a fuller expression is preferred to refer to it, for instance, the overt form
- The results for the SVO+overt condition are in line with results obtained by Costa and colleagues (1998, 2004) and Carminati (2002), for instance, that found that the overt form of the pronoun is more agnostic than the covert one. Although preferentially chosen to retrieve the OBJECT, it is also sometimes chosen to refer to the SUBJECT, even in conditions where the SUBJECT is also the first mentioned entity
- Our results also suggest that the identification of the null and the overt pronoun occurs at different moments, being the identification of the overt pronoun sooner than its null counterpart. On the null conditions we only find differences between looks to Subject and looks to Object 600ms after the onset of the pronoun/verb, contrary to what we find in the overt conditions, where differences emerge at the first period, between 200ms and 400ms after the onset of the pronoun
 - A spillover effect due to recency cannot be excluded, although not present in all conditions. Also, as it may be observed, even though at the 200ms-400ms interval it may be the case (in OVS+NULL or SVO+OVERT), at 400ms-600ms there is little or no difference between looks to SUBJECT and looks to OBJECT in SVO+NULL, OVS+NULL and SVO+OVERT condition, while there are, clearly, more looks to the SUBJECT in the OVS+OVERT condition

5. Conclusions

Our results suggest that different types of syntactic information, like Subjecthood and Topicality, contribute to saliency ascription differently, depending on if they are combined or pitted against each other. An antecedent that is both the Subject and the Topic of the sentence (first-mentioned entity) will be clearly preferred to co-refer with a null pronoun equally during sentence processing and at a later stage (results from SVO+NULL condition). However, if those sources of information are distributed among different entities, results from on-line and off-line tasks may differ, showing some conflict in the parser (results from OVS+NULL condition). Even if the same results are found in off-line and on-line tasks, they will be different from results from conditions with different combination of factors (SVO+NULL and OVS+OVERT), when the two factors are combined the results are stronger than when there is a conflict between them.

As predicted, our Third Hypothesis is confirmed by the results: prominence is the result of the combination of various factors, namely Syntactic-Function and Order of Mention. That is, as proposed by Kaiser (2006:151), saliency is not a monolithic concept.

References

- Alonso-Ovalle, L., Clifton, C., Frazier, L., Solera, S. (2002). Null vs. Overt Pronouns and The Topic-Focus Articulation, in Spanish. *Journal of Italian Linguistics*, 14, 151-169
- Ariel, M. (1996). Referring expressions and the +/- coreference distinction. In J. Gundel & Fretheim (eds.), *Referent and Referent accessibility* (pp. 13-35). John Benjamins
- Arnold, J., Eisenband, J., Brown-Schmidt, S., Trueswell, J. (2000). The immediate use of gender information: eyetracking evidence of the time-course of pronoun resolution. *Cognition*, 76, 813-826
- Carminati, M. (2002). The processing of Italian subject pronouns. Doctoral dissertation, University of Massachusetts at Amherst
- Carreiras, M., Gernsbacher, M., Villas, V. (1995). The advantage of first mention in Spanish. *Psychonomic Bulletin and Review*, 2, 124-129
- Corrêa, L. (1998). Acessibilidade, paralelismo na interpretação do pronome sujeito e o contraste pro/pronome em português. *DELTA*, 24, 295-392
- Costa, A., Faria, I. H., Matos, G. (1998). Ambigüidade referencial na identificação do sujeito em estruturas coordenadas. *Actas do XIII Encontro Nacional da Associação Portuguesa de Linguística*. Lisboa: APL, 173-188
- Costa, A., Faria, I., Kail, M. (2004). Semantic and Syntactic Cues' Interaction on Pronoun Resolution in European Portuguese. In Branco, McEnery, & Mitkov (Ed.), *DAARC 2004, 5th Discourse Anaphora Resolution Colloquium*. Lisboa: Colibri, 45-50
- Gernsbacher, M., Hargreaves, D. (1988). Accessing sentence participants: The advantage of first mention. *Journal of Memory & Language*, 27, 699-717.
- Järvikivi, J., Gompel, R., Hyönä, J., Bertram, R. (2005). Ambiguous pronoun resolution: Contrasting the first-mention and subject-preference accounts. *Psychological Science*, 16(4), 260-264
- Kaiser, E., Trueswell, J. (2008). Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference. *Language and Cognitive Processes*, 23(5), 709-748
- Kaiser, E., Vihman, V. (2006). On the referential properties of Estonian pronouns and demonstratives. In Hans Götzsche (ed.), *Memory, Mind and Language*. Cambridge Scholars Publishing, Newcastle, 193-205
- Kaiser, E. (2006). Effects of topic and focus on saliency. In Christian Ebert & Cornelia Endriss (eds.), *Proceedings of Sinn und Bedeutung 10*, ZAS Working Papers in Linguistics, vol. 44, 139-154, Berlin
- Melo, M., Maia, M. (2005). O processamento da co-referência do sujeito pronominal em sentenças formadas por verbos de comunicação linguística no português do Brasil. *Linguística*, 1, 177-206

This study was funded by the FCT Project: PTDC/LIN/67854/2006



CLUL Centro de Linguística da Universidade de Lisboa

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR