The production of null and overt subject pronouns in Italian and Spanish: Effects of Semantic Role Predictability

The Position of Antecedent Hypothesis was proposed by Carminati (2002) to explain the interpretation of null and overt pronouns in Italian (a null-subject language). The Position of Antecedent Hypothesis suggested that pronoun interpretation is influenced by the structural position of its antecedent within a sentence: While null pronouns are interpreted towards antecedents in a prominent syntactic position (subject), overt pronouns prefer antecedents in lower positions. On the other hand, according to statistical learning theories, individuals learn to produce and interpret pronouns by observing patterns of usage, making predictions based on context, and using frequency information (Arnold, 2001). For example, speakers are not only sensitive to syntactic cues like subjecthood when interpreting pronouns, but they also have expectations about how semantic patterns influence the interpretation of pronouns.

In the present study, we focus on Mexican Spanish and Italian, two null-subject languages where both null and overt pronouns are used. A study by Medina et al. (2022) on Spanish using source and goal verbs, showed that when the semantic role of the referent is highly predictable in a given context, speakers are less likely to use an overt pronoun. In contrast, when the semantic role is less predictable, speakers tend to use an overt pronoun to clarify the reference. Here, we investigate further how semantic roles and discourse expectations interact to shape pronoun use in null-subject languages, by looking at Italian and Mexican Spanish. We analyze production biases in the two languages, including a comparison of references to subject/object antecedents in contexts of intra-sentential anaphora, where two referents sharing similar gender are introduced, and implicit causality verbs are included. In contexts with implicit causality verbs, thematic role predictability should bias speakers to produce a sentence continuation where either the preceding subject or the preceding object is mentioned using a null/overt pronoun.

A written sentence completion task was created where sentence fragments contained two antecedents with similar gender (e.g., Mike and Brian), one in subject position and one in object position. The fragment included a subject-biased (Mike *frightened* Brian because...) or an object-biased implicit causality verb (Mike *blamed* Brian because...). Sentence completions were classified according to the type of referential expression produced (null/overt pronoun and proper noun; the dependent variable) and according to the referent of the referential form (referring to the preceding subject or object; independent variable: Antecedent type). Thirty-five* Italian native speakers and seventy-six Mexican Spanish native speakers participated in the written sentence completion task.

The preliminary results show that Italian and Spanish-speaking participants produced null pronouns at ceiling to refer to a subject or an object antecedent (Table 1). The statistical analysis showed that significantly more null pronouns were produced to refer to subject antecedents than object antecedents (main effects of Antecedent, all ps<.01). In addition, both groups of speakers produced fewer overt pronouns to refer to subject antecedents than to refer to object antecedents (main effects of Antecedent, all ps<.01).

The preliminary results show that in Mexican Spanish and Italian, pronoun use is tightly coupled with predictability of thematic roles, and predictable referents are strongly associated with null pronoun production. Based on our findings, we propose that predictability in discourse is not solely determined by syntactic prominence as suggested by the Position of Antecedent Hypothesis (Carminati, 2002), but also by the semantic role that the referents play in the events described, as demonstrated by the implicit causality context in our study. Thus, models of pronoun production need to account for semantic role predictability as a significant factor in the use of pronouns in null-subject languages like Italian and Mexican Spanish (Medina et al., 2022).

* Data collection is currently ongoing: Italian-speaking participants are currently recruited to match the number of Mexican Spanish speakers.

Table 1. Proportion of intra-sentential null/overt pronouns and proper nouns produced by Italian (N=35) and Spanish (N=76) speakers in reference to a preceding subject and object referent.

Subject-reference (Mike *frightened* Brian because….)	Null Pronoun	Overt Pronoun	Proper Noun
Italian	0.99	0.01	0.00
Spanish	0.95	0.04	0.02
Object-reference (Mike *blamed* Brian because…)			
Italian	1	0	0
Spanish	0.85	0.10	0.05

References

Arnold, J. E. (2001). The effect of thematic roles on pronoun use and frequency of reference continuation. *Discourse Processes, 31*(2), 137–162.

Carminati, M. N. (2002). *The processing of Italian subject pronouns* [Doctoral Dissertation], University of Massachusetts at Amherst.

Medina Fetterman, A. M., Vazquez, N. N. & Arnold, J. E. (2022). The effects of semantic role predictability on the production of overt pronouns in Spanish. *Journal of Psycholinguistic Research*, 1–26.