

Attachment Preferences and Corpus Frequencies in Prepositional Phrase Structures: A Study in Greek

Kalliopi Katsika

Aristotle University of Thessaloniki

kkatsika@enl.auth.gr

1 Introduction

The present study investigates native Greek speakers' online parsing preferences for ambiguous V-NP-PP structures and compares them to structural frequency data from a corpus analysis in Greek. The comparison between these two types of linguistic evidence was conducted in order to examine the assumptions of the Tuning Hypothesis (Cuetos, Mitchell, and Corley, 1996) which directly links parsing preferences to structural frequency, as it assumes that the resolution of ambiguous structures is determined by the reader's or listener's prior experience / exposure to ambiguities of the same kind. The extent to which the Tuning Hypothesis can actually account for parsing preferences has been under investigation in the psycholinguistic literature.

Previous findings on the possible relevance between corpus frequencies and parsing preferences have been controversial. Gibson and Pearlmutter (1994) conducted a corpus analysis which involved structures of the form "*NP₁ Prep NP₂ Prep NP₃ Prep NP₄*". The extracted corpus structures were compared to previous psycholinguistic data and the partial compatibility of the two types of linguistic evidence was claimed by the authors to cast "... doubt on the hypothesis that corpora and comprehension data are completely irrelevant". Nevertheless, in more recent work on noun phrase conjunctions, Gibson and Schütze (1999) found a discrepancy between corpus frequencies and attachment preferences. Moreover, cross-linguistic evidence from findings in Dutch by Desmet, Brysbaert, and De Baecke (2002) and the study by Desmet and Gibson (2003) reveal that structural frequencies may reflect comprehension preferences when the corpus data are carefully controlled in order to match the experimental sentences of the comprehension tasks.

In the present study, native Greek speakers' PP attachment preferences were investigated on the basis of two types of linguistic evidence; corpus data and comprehension preferences. Online self-paced reading task results were thus compared to data from a sample text analysis of the 46,000,000 word Greek corpus developed by ILSP (Institute for Language and Speech Processing).

2 PP Attachment Ambiguity

Prepositional Phrase attachment ambiguities involve structures such as (1) which are ambiguous in that the PP *me to tileskopio* can be attached either to the verb *idhe* denoting the instrument of the action described by the verb or to the preceding NP *ton andhra* as a modifier:

- (1) *O kataskopos idhe ton andhra me to tileskopio.*
The-NOM spy-NOM saw-PERF.3S the-ACC man-ACC with the-ACC telescope-ACC
'The spy saw the man with the telescope.'

It has been claimed in the psycholinguistic literature that native speakers' initial attachment preferences may be based upon different types of information such as syntactic (Frazier, 1987; Rayner, Carlson, and Frazier, 1983), subcategorization (e.g. Clifton, Speer, and Abney, 1991), discourse level / lexical information (e.g. MacDonald, Perlmutter, and Seidenberg, 1994) or structural frequencies (Cuetos et al., 1996).

The present study involves the investigation of native Greek speakers' initial attachment preferences for temporarily ambiguous V-NP-PP structures through an online self-paced reading task. The experimental stimuli included sentences in which four Greek prepositions (*me* 'with', *se* 'in, into', *ja* 'for', *apo* 'from, by') were inserted. The experimental sentences thus had the following form:

- (2) a. *O kipuros ekopse to kladhi me to maheri.* (VP Attachment)
The-NOM gardener-NOM cut-PERF.3S the-ACC branch-ACC with the-ACC knife-ACC
'The gardener cut the branch with the knife.'
- b. *O kipuros ekopse to kladhi me to luludhi.* (NP Attachment)
The-NOM gardener-NOM cut-PERF.3S the-ACC branch-ACC with the-ACC flower-ACC
'The gardener cut the branch with the flower.'

The results indicated a significant main effect of attachment during the processing of the critical PP segment both in the subjects and in the item analysis; $F_1(1,46) = 13.293, p < .001, \eta^2 = .224, F_2(1,31) = 12.647, p < .001, \eta^2 = .290$. VP attachment was significantly preferred over NP attachment in the case of prepositions *me* ($t(46) = 2.438, p < .02$), *ja* ($t(46) = 3.627, p < .001$) and *apo* ($t(46) = 3.712, p < .001$), whereas preposition *se* did not induce any significant effect of attachment on the critical segment. These results were partially compatible with the corpus data in which VP attachments were more frequent than NP attachments in the case of prepositions *me*, *se* and *apo*. In the case of preposition *ja*, however, NP attachment structures were significantly more frequent than VP attachment structures.

3 Conclusion

The comparison between the corpus data and the online self-paced comprehension results indicated a partial correspondence between the two types of linguistic evidence. Thus, on the basis of the results of the present study, one could not make strong claims regarding the reflection of corpus frequencies to parsing preferences. Possible further filtering of the corpus data may provide a clearer picture as to whether or not parsing preferences in Greek actually correspond to corpus frequencies.

References

- Clifton, C., Speer, S., and Abney, S.P. (1991). Parsing Arguments: Phrase Structure and Argument Structure as Determinants of Initial Parsing Decisions. *Journal of Memory and Language*, 30: 251-271.
- Cuetos, F., Mitchell, D. C., and Corley, M. M. B. (1996). Parsing in different languages. In M. Carreiras, J. Garcia-Albea, & N. Sabastian-Galles, eds., *Language Processing in Spanish*. Hillsdale, NJ: Erlbaum.
- Desmet, T., M. Brysbaert, and C. De Baecke (2002). The correspondence between sentence production and corpus frequencies in modifier attachment. *The Quarterly Journal of Experimental Psychology*, 55A(3): 879–896.
- Desmet, T., M. and Gibson, E. (2003). Disambiguation preferences and corpus frequencies in noun phrase conjunction. *Journal of Memory and Language*, 49: 353-374
- Frazier, L. (1987). *Syntactic Processing: Evidence from Dutch*. *Natural Language and Linguistic Theory*, 5: 519-559.
- Gibson, E. and Pearlmutter, N.J. (1994). A corpus-based analysis of psycholinguistic constraints on Prepositional-Phrase attachment. In C. Clifton, L. Frazier and K. Rayner, eds, *Perspectives on Sentence Processing*, Hillsdale, N.J.: Lawrence Erlbaum: pp. 181-198
- Gibson, E. and C. T. Schütze (1999). Disambiguation preferences in noun phrase conjunction do not mirror corpus frequency. *Journal of Memory and Language*, 40: 263–279.
- MacDonnald, M.C., Pearlmutter, N.J., and Seidenberg, M.S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101: 676–703.
- Rayner, K., Carlson, M., and Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior*, 22: 358-374.