**The Processing of Italian Relative Clauses within a Computational Parsing Model**

Aniello De Santo, Stony Brook University

**Overview** Restrictive relative clauses in Italian have been the focus of extensive experimental studies from the perspective of comprehension, production, and acquisition. Apart from conforming to a well-attested cross-linguistic preference for subject over object relative clauses, Italian speakers also show increased processing difficulties when encountering relative clauses with subjects in postverbal position. We show how these processing differences can be accounted for via a computational model that relies on "memory burden" to connect predictive, incremental parsing to specific syntactic analyses (Graf et al. 2017:a.o.), and discuss the relevance of this model for existing theories of sentence processing.

**Asymmetries in Italian Relative Clauses** Italian speakers show a general preference for subject over object relative clauses, so that (1) is easier to process than (2) (Volpato and Adani 2009:a.o.):

(1) Il cavallo che insegue i leoni **SRC**
"The horse that chases the lions"

(2) Il cavallo che i leoni inseguono **ORC**
"The horse that the lions chase"

Interestingly, Italian also allows for object relatives like (3), where the embedded subject is expressed postverbally and a null *pro* is postulated in the embedded preverbal position.

(3) Il cavallo che *pro* inseguono i leoni **ORCp**
The horse that *pro* chase the lions
"The horse that the lions chase"

Although postverbal subject constructions are very common in Italian, studies comparing the processing of distinct kind of RCs have reported increased efforts with ORCp, leading to the following difficulty gradient: $SRC < ORC < ORCp$ (Utzeri 2007:a.o.).

**Minimalist Grammar Parsing and Memory Usage** We are interested in formulating a quantifiable theory of the effects of grammatical structure on sentence processing. This is the advantage of the computational model adopted here, in that it provides a quantitative measure of the way subtle structural differences affect memory resources. Recently, several studies have shown how Stabler (2013)'s top-down parser for Minimalist grammars (MGs) can be combined with complexity metrics that relate parsing difficulty to memory usage, and successfully used to explain processing difficulty across a variety of constructions (Kobele et al. 2013:a.o.). In particular, the MG parser refers to three cognitive notions of memory usage: I) how long a node is kept in memory (TENURE); II) how many nodes must be kept in memory (PAYLOAD); or III) how many bits a node consumes in memory (SIZE). Based on these abstract, psychologically grounded concepts, Graf et al. (2017) define a set of metrics measuring processing difficulty, which we use here to analyze the Italian data.

**Modeling Italian RCs** We test the parser performance on sentences of the form *I saw the horse* $[_{RC}$ *that ...]*, with the embedded relative clause either an SRC (1), an ORC (2), or an ORCp (3). The choice of syntactic analysis is particularly important, due to the parser's sensitivity to grammatical structure. Our analysis of postverbal subjects (sketch in 4) follows Belletti and Contemori (2009).

(4) Il cavallo$_i$ che $[_{TP}$ *pro* inseguono $[_{VP}$ <V> <cavallo$_i$>] … i leoni$_j$ … $[_{vP}$ <i leoni$_j$> <VP>]]

In ORCp constructions, the subject DP *[i leoni]* is merged in preverbal subject position Spec,*v*P, and then raised to a Spec,Focus position in the clause-internal *v*P periphery; the whole verbal cluster is

raised to a clause-internal Spec,Topic position; and an expletive *pro* is base generated in Spec,TP. Moreover, in line with most of the psycholinguistic literature on this topic, we adopt a promotion analysis of relative clauses (cf. Arosio et al. 2017:a.o.). That is to say, the head noun starts out as an argument of the embedded verb and undergoes movement into the specifier of the relative clause.

Our simulations show that the parser correctly predicts the gradient of difficulty observed for Italian RCs ($1 < 2 < 3$), across a variety of memory metrics. In particular, the MAXIMUM TENURE measured over the trees turns out to play a crucial role in discriminating between alternatives, reflecting the burden of resolving movement dependencies during processing. Since MAX. TENURE has been noted as an elegant measure of processing difficulty in previous studies (Graf et al. 2017), this result also supports the plausibility of the MG parser as a general model of sentence processing.

**Discussion** Importantly, the way high TENURE values are driving the parser' preferences in our simulations shows that the additional movement dependencies postulated for postverbal subject constructions play a crucial role in the increased processing efforts for ORCp.
Obviously, the model adopted here is not the first associating some kind of memory cost to long-distance dependencies. What we claim is that a precise specification of the parsing model allows us to reinterpret previous theories in a quantifiable framework that directly connects parsing processes to cognitive resources. Consider De Vincenzi (1991)'s Minimal Chain Principle (MCP), which is commonly referred to as a way to ground Italian RC asymmetries in parsing effects. The MCP postulates that shorter dependencies are computationally less demanding than longer dependencies: thus SRCs are easier than ORCs because the filler gap distance in the former is shorter than in the latter. *Economy principles* also predict the increased difficulty found for ORCs with postverbal subject, since these involve two chains instead than one. However, it is not clear how these *computational demands* would be implemented in a precise parsing architecture, and how these costs are linked to cognitive resources like working memory, known to affect processing effects (Utzeri 2007). The MG parser does not directly postulate chains, but it explicitly connects processing differences to the additional memory resources involved in keeping track of long movement dependencies. Thus, it offers a way to reinterpret De Vincenzi's theory, and more general filler-gap dependencies results, in a framework that takes both economy claims and syntactic assumptions seriously.

The MG parser is also compatible with a line of research which attempts to reduce efforts in ORC processing to interference effects caused by the relative head moving across an embedded subject endowed with a similar feature set (*featural Relativized Minimality*; Arosio et al. 2017:i.a.). This approach does not fully predict differences in performance on ORC over ORCp, but the evidence for interference effects modulating the processing difficulty of RCs is compelling. Importantly, while it is unclear how the MCP could incorporate these effects, the MG model offers direct ways to accommodate them, since the parser is grounded in a feature-driven formalism. In the future, the transparent specification of the parser' behavior will allow us to clearly quantify the contribution of different components (length of dependencies, feature overlap, etc.) to sentence processing effects.

**Arosio, F. et al.** (2017). The comprehension of Italian relative clauses in poor readers and in children with specific language impairment. *Glossa*. • **Belletti, A. and Contemori C.** (2009). Intervention and attraction. on the production of subject and object relatives by Italian (young) children and adults. In *Proceedings of Gala*. • **De Vincenzi, M**. (1991). *Syntactic parsing strategies in Italian: The minimal chain principle*. Springer Science & Business Media. • **Graf, T. et al.** Relative Clauses as a Benchmark for Minimalist Parsing. *Journal of Language Modelling*. • **Kobele, G.M. et al.** (2013). Memory resource allocation in top-down minimalist parsing. In *Formal Grammar*. Springer. • **Stabler, E.P.** (2013). Bayesian, minimalist, incremental syntactic analysis. *Topics in Cog. Sci.* 5:611–633. • **Utzeri, I.** (2007). The production and acquisition of subject and object relative clauses in Italian: a comparative experimental study. *Nanzan Linguistics* 2. • **Volpato, F. and Adani F.** (2009). The subject/object relative clause asymmetry in Italian hearing-impaired children: evidence from a comprehension task. *Studies in Linguistics* 3.