

Testing accessibility: A cross-linguistic comparison of the syntax of referring expressions

Jacopo Torregrossa, Christiane Bongartz, Ianthi Tsimpli¹

LSA Annual Meeting 2015, Portland 8-11 January 2015

1. Introduction. Focusing on the discourse conditions that license the use of null subjects (*pro*) in Greek and Italian, this paper shows that the distribution of referring expressions (RE, e.g., overt and null pronoun, clitic, definite description, etc.) does not only depend on the referents' discourse status (alias accessibility). Syntactic constraints play an important role too.

According to Ariel's Accessibility Theory (Ariel 1990 and following work), the form of a RE encodes the degree of accessibility of its discourse antecedent. More specifically, REs are ordered on a scale, ranging from low to high accessibility markers. For instance, if compared to pronouns (e.g., *he*), definite descriptions (e.g., *the man*) are more informative, more rigid (i.e., able to refer to a unique referent) and phonologically heavier, and express a lower degree of accessibility of their discourse antecedents. (1)-(3) are exemplifications of how different types of REs are ordered in the accessibility marking scale in English, Greek and Italian respectively. While, according to Ariel, the existence of a scale of REs is a universal feature of languages, the inventory of REs is language specific. For instance, English has both stressed (*HE*) and unstressed (*he*) pronouns, and does not have clitics and *pros*. On the contrary, in Greek and Italian destressing of pronouns is not allowed, while clitics and *pros* exist and are in complementary distribution.²

- (1) English: *the old man* < *the man* < *that man* < *this man* < *HE* < *he* < \emptyset (e.g., gaps)
- (2) Greek: *ο γέρος άνθρωπος* < *ο άνθρωπος* < *εκείνος ο άνθρωπος* < *αυτός ο άνθρωπος* < *αυτός* < *τον*_{ACC-MASC}, *pro*
- (3) Italian: *l'uomo anziano* < *l'uomo* < *quell'uomo* < *quest'uomo* < *lui* < *lo*_{ACC-MASC}, *pro*

If the use of a RE were only determined by the accessibility of the discourse antecedent, to a certain accessibility value would correspond the use of the same type of RE in Greek and Italian, given that the two languages split the accessibility hierarchy along the same categories (as shown by (2) and (3)).

According to the relevant literature, the antecedent's degree of accessibility is sensitive to different linguistic factors, such as its grammatical role, syntactic position and the recency of its last mention (see Arnold 2010 for a review). For instance, subject referents (e.g., *John* in (4)) are more accessible than non-subject referents (*Charles*), and last mentioned referents are *ceteris paribus* the most accessible (see (5)). This explains why *John* in (4) and *Charles* in (5) can be picked up by a RE encoding high accessibility, i.e., the unstressed pronoun *he*.

¹ Jacopo Torregrossa, University of Cologne (jtorregr@uni-koeln.de), Christiane Bongartz, University of Cologne (teambongartz@googlemail.com) & Ianthi Tsimpli, University of Thessaloniki and University of Reading (imt@enl.auth.gr).

² Ariel argues that clitics encode a lower degree of accessibility than *pros*. This assumption cannot be verified in Italian and Greek, given their complementary distribution. Therefore, these two categories are placed at the same level in the scales in (2) and (3). As for English, only first person singular subjects can be omitted (see, a.o., Haegeman 2013).

- (4) John_i went jogging with Charles_j. He_{i/*j} ran for two hours.
 (5) John_i ate a cake. Charles_j drank some milk. He_{*i/j} asked the bill.

In this paper, we investigate whether the antecedent's degree of accessibility can account for the distribution of REs in Greek and Italian, with particular reference to *pros*. (6) – adapted from Frascarelli (2007:725) – shows that in Italian *pro* cannot refer to the preceding indirect object *a Mario* (which in the example is selected by the verb *raccontare* 'tell' in the relative clause contained in the clitic left dislocated constituent). However, coreference between the two constituents is accepted in Greek, as shown in (7).

- (6) * [La storia che hanno raccontato a Mario]_k_i, la_i *pro*_k dice a tutti.
 the story that AUX-3PP told to M. CL-ACC-FM tell-3SG to everyone.
 (7) [την ιστορία που ειπαν στη Μαιρη]_k_i τη_i *pro*_k λει σε όλους.
 the story that tell-PST-3PP to M. CL-ACC-FM tell-3SG to everyone.
 'The story that they told to Mario, he tells to everybody'.

(6) and (7) suggest that *pro* encodes different degrees of accessibility of its antecedent in the two languages. This pattern is expected under the syntactic analysis of *pro* which has been proposed in the literature for Italian and Greek respectively. According to Frascarelli (2007), Italian *pro* sits at the edge of *vP* and is licensed by an Aboutness Topic in the CP, as indicated in (8). Crucially, the constituent *a Mario* is not visible at the edge of the phase containing it, and the formation of the chain *Mario-a Mario-pro* is blocked. This renders coreference between *a Mario* and *pro* impossible.³

- (8) <Mario>_k [la storia che hanno raccontato a Mario]_k_i, la_i *pro*_k dice a tutti.

Greek *pro* is not licensed in the same way as Italian. Christodoulou (2007) accounts for coreference relations between Greek pronouns and their antecedents in terms of linear precedence (as shown in (7)). More in general, Roussou & Tsimpli (2006) argue that Greek subjects are licensed IP-internally in the specifier of TP, contrary to what happens in Italian.

This paper assesses the degree of accessibility encoded by each type of RE in Greek and Italian, based on a multifactorial analysis of REs produced in the context of a story-telling task by adult speakers in both languages. In the light of the previous discussion, two scenarios are possible: i) REs encode the same accessibility values in both languages; ii) the same type of RE encodes different accessibility values in the two languages. The results show that ii) holds true. This difference is consistent with the syntactic account.

³ This paper is mainly concerned with topics that are subjects – we follow Frascarelli (2007) in arguing that overt subjects in Italian occupy a TopP in the CP. However, (i) – taken from Frascarelli (2007:726) shows that *pro* can corefer with other types of arguments. For example, coreference of *pro* with the dative clitic *gli* in (i) is accepted. This is compatible with Frascarelli's theory, since clitics are licensed by a topic (overt or null) in the CP: Chain-formation between the left dislocated indirect object (*a Mario*), the clitic (*gli*) and *pro* is thus possible.

(i) <A Mario>_i [la storia che gli_i hanno raccontato]_k, la_k *pro*_i dice a tutti.
 To Mario the story that CL-DAT-MASC AUX-3PP told it-CL-ACC-FEMM tell-3SG to everybody
 Note that the configuration in (i), with *pro* having a clitic as antecedent, accounts for the (few) instances of Italian *pros* preceded by a non-subject antecedent reported in Figure 2 under Results. On the contrary, many of the non-subject arguments preceding *pros* in Greek are DPs in IP-internal position.

2. Methods. Participants. 20 Italian (15 females and 5 males; age range: 20-38) and 16 (15 females and 1 male; age range: 20-40) Greek adult speakers took part in the study. The Italian participants were recruited at the University of Macerata (Italy), while the Greek participants at the University of Thessaloniki (Greece). All participants were monolinguals and had no history of language impairment. **Materials.** The production of REs was analyzed in the context of a story telling task. Each participant produced two narratives, elicited by means of the picture stories from the Edmond Narrative Norms Instrument (ENNI; Schneider et al. 2006: http://www.rehabmed.ualberta.ca/spa/enni/about_the_enni.htm): the first narrative consisted of 13 pictures and involved two main and two minor characters, while the second narrative consisted of 8 pictures and involved two main characters and one minor character. Narratives were told to a monolingual investigator, who feigned ignorance of the stories. **Analysis.** Unit of the analysis was the clause defined by the occurrence of a verb. We analyzed only REs denoting animate characters, coding them for type (null, clitic, DP), for their features (clause – indicating whether they appeared in a *main* or a *subordinate* clause – and argument status – *subject*, *object*, *other*), for features of the antecedent (a-clause and a-argument) and for distance from the antecedent (measured in number of clauses). Table 1 presents an extract from the analysis of Italian data. A decision tree analysis was carried out to identify which of these features plays a major role in determining the use of REs.

UNIT	ENGLISH TRANSLATION	R-type	R-clause	R-argument	A-type	A-clause	A-argument	distance
E <i>l'elefante</i> fa un bagno in piscina	And <i>the elephant</i> has a bath in the swimming pool	DP	MAIN	SUBJ	DP	MAIN	OTHER	3
Però non <i>Ø</i> ha letto il cartello.	But <i>he</i> has not read the sign	NULL	MAIN	SUBJ	DP	MAIN	SUBJ	1

Table 1: A sample of the analysis.

3. Results and discussion. The decision tree in Figure 1 singles out the most relevant factors influencing the use of REs in Italian. The features are ordered in a top-down manner, from the strongest to the weakest predictor of the RE-type.

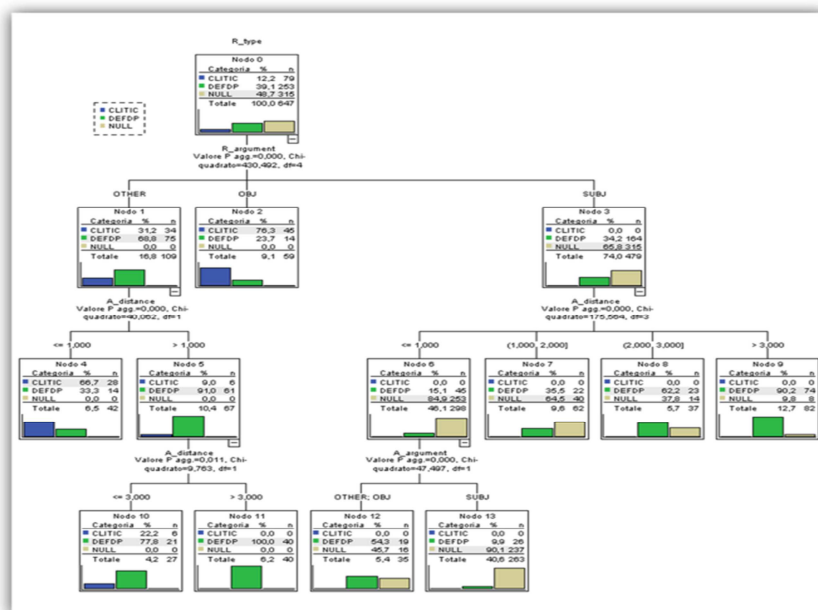


Figure 1: Decision tree of the discourse factors that affect the use of REs in Italian.

The RE's argument status is the root node (i.e., the best predictor of RE). This is expected, given that *pros* are necessarily subjects, while clitics non-subjects. Distance of the RE from its antecedent is the next strongest factor: high accessibility markers (i.e., *pros* and clitics) tend to be used at shorter distances. However, whenever the distance is less than or equal 1, it is the antecedent's argument status that affects the RE type: If the antecedent is a subject, the RE is consistently a null. Otherwise, it tends to be a DP.

Crucially, the decision tree concerning the use of Greek REs – which is not shown here for reasons of space – patterns the same as far as the first two predictors (RE's argument and distance) are concerned. However, the two languages differ in the nature of the last decision node. The factor 'antecedent's argument' in Greek is not as influencing as it is in Italian, namely, the distribution of *pros* in Greek is less constrained by the antecedent's argument status than in Italian. Instead, the third factor that influences the use of REs in Greek is the antecedent's referential type. Figure 2 shows that the percentage of NPs that are preceded by a non-subject antecedent at different distances (i.e., 1 and equal to or more than 2) in Greek is significantly greater than in Italian ($\chi^2(1)=3.96, p<.05$). This confirms the pattern emerging from the decision tree analysis.

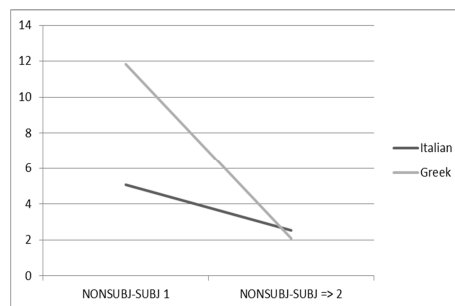


Figure 2: Percentage of *pros* whose antecedent is not a subject at different distances in Greek and Italian.

The multi-factorial analysis carried out in this paper shows that the discourse conditions that affect the use of REs in Greek and Italian are similar in many respects (e.g., argument status of the RE, distance from the antecedent, etc.), as predicted by the similar way in which the two languages split the Accessibility Marking Scale (see (2)-(3) above). However, a difference can be observed in the argument status of the antecedent of *pros*: In Italian the antecedent must be a subject (or more in general a left dislocated constituent – see footnote 3), while this is not the case for Greek. This pattern of variation supports the syntactic analysis of *pro* presented in Section 1.

References

- Ariel, Mira. 1990. *Accessing noun-phrase antecedents*. London: Routledge.
- Arnold, Jennifer. 2010. How speakers refer: The role of accessibility. *Language and Linguistic Compass* 4:4. 187-203.
- Christodoulou, Christiana. 2007. Anaphoric relations with Greek pronouns. *Proceedings of the North West Linguistics Conference*. Victoria B.C.
- Frascarelli, Mara. 2007. Subjects, topics and the interpretation of referential *pro*: An interface approach to the linking of (null) pronouns. *Natural Language and Linguistic Theory* 25:4. 691-734.
- Haegeman, Liliane. 2013. The syntax of registers: Diary object omission and the privilege of the root. *Lingua* 130. 88-110.
- Roussou, Anna & Ianthi Maria Tsimpli. 2006. On Greek VSO again!. *Journal of Linguistics* 42. 317-354.