Thematic roles affect pronoun production in Romanian

Sofiana I. Lindemann – Universitatea Transilvania din Brasov
Sofiana.chiriacescu@unitbv.ro

1 Introduction

Research has convincingly shown that pronoun resolution is influenced by a range of factors on the sentence as well as on the discourse level. On the sentence level, one of the factors that seems to play a central role is grammatical function. A referent realized as the sentence subject is more prone to be subsequently pronominalized compared to a non-subject referent (Givón 1981, Ariel 1990, Gundel, Hedberg & Zacharski 1993, Grosz, Joshi & Weinstein 1995, Chiriacescu 2011). The picture is less clear with respect to the role played by different thematic roles in guiding pronoun processing. While one line of research showed that thematic roles do not affect reference form production (Kehler, Kertz, Rohde & Elman 2008, Fukumura & van Gompel 2010), other studies argued that thematic roles influence pronoun usage (Arnold 2001, Kaiser, Li & Holsinger 2011, Rosa & Arnold 2017). The present study investigates these two options, by exploring the impact of different thematic roles on pronoun production in Romanian.

Romanian is a pro-drop language (Dobrovie-Sorin 1994), which allows for both overt and null pronouns in preverbal subject position. The widespread view regarding its pronominal system is that null pronouns are predominantly used for the most accessible antecedent, while overt pronouns are employed to refer back to a less accessible antecedent, with accessibility being primarily driven by grammatical function. Accordingly, null pronouns are typically referring back to preverbal subjects and overt pronouns are generally used for a lower syntactic position (see e.g. Zafiu 2008, Pagurschi 2010, Teodorescu 2016, on Romanian and Carminati 2002 for Italian).

This is the first study exploring whether a semantic factor affects reference form in Romanian language production. The main goal is to determine whether the alternation between null and overt pronoun production is guided by different thematic roles, in addition to the long attested subjecthood constraint. If thematic roles do play a role as well, we expect this effect to add up to the grammatical function bias. Second, the findings will bring further evidence to the debate whether thematic roles affect reference form production or not.

2 The experimental study

Participants

50 native speakers of Romanian from the Transilvania University of Brasov, Romania, took part in the experimental study (age range: 18–45 years, mean age 28 years, 29 female). It took about twenty minutes to complete each version of the study.

Design, procedure and materials

A written story completion task was used in which each target sentence consisted of two human referents that had the same gender (e.g. Mihai and Anton in the first example in Table 1). Speakers tend to prefer less explicit types of referring expressions (e.g. pronouns) to refer back to a previously introduced subject referent, thus, in order to keep grammatical and thematic roles apart, we used transfer-of-possession verbs. These verbs are a good testing case, as some of them realize the Goal as the grammatical subject (e.g. *get* in CND_2 in

Table 1), while others realize the Source in grammatical subject position (e.g. *give* in CND_1 in Table 1). Participants' task consisted in reading the given one-sentence target items and providing one natural sounding written sentence continuation to each item. We crossed thematic roles (Goal vs. Source) and grammatical function (Subject vs. Object) and used 18 experimental items, 9 for each condition, and 20 filler items, distributed in two lists. In line with the findings from previous studies on pronoun production, the predictions are that null pronouns will prefer a subject antecedent, whereas overt pronouns will preferentially pick up the non-subject referent. If the manipulation of thematic role matters as well, the expectation is that Goal referents will increase the rate of both null and overt pronominalization, irrespective of grammatical function (Rosa & Arnold 2017).

Table 1: Example items with English translations

| CND1_Goal=Non-Subject | Mihai i-a dat un bilet lui Anton. Mihai gave a note to Anton. |
| CND2_Goal=Subject | Cristian a primit o carte de la Dan. Cristian got a book from Dan. |

Results

Two independent coders manually annotated the type of referring expression (i.e. null pronoun, overt pronouns, clitic, demonstrative pronoun, proper name, definite noun phrase, etc.) chosen by the participants to refer back to the referent. They coded 900 continuations.

The general finding is that both grammatical function and grammatical role matter. Overall, given all overt and null pronouns produced by participants in their continuations, the majority was used to pick up the subject referents (72%). Furthermore, the most robust finding is that more null pronouns were used to pick up the previous subject referent (77%), irrespective of its thematic role. This is in line with previous studies and the predictions made (Ariel 1990, Zafiu 2008).

The overt pronoun seems to be more versatile, as it is used to refer back to either the previous subject or non-subject referent, however, still with a slight preference for the previous subject (57%) rather than the non-subject (43%). The findings indicate that we do not find a clear division of labour between the two pronoun forms in terms of syntactic structure, as was expected based on the existing literature. The observed biases can be best accounted for in terms of preferences: the null pronoun has a more robust preference towards the subject antecedent than does the overt pronominal form.

Interestingly, we found an additional effect of thematic role. More concretely, the realisation of a referent as the thematic Goal (vs. the Source) contributed to its predictability in terms of next mention and pronominalization. Both subject and non-subject referents benefited from the thematic role manipulation, such that the overall pronominalization rates increased for the Goal referents (i.e. the non-subject in CND1 and the subject in CND2). Being realized as the thematic Goal particularly impacted overt pronoun production. In CND2, 93% of the personal pronouns were used for the subject=Goal antecedent (vs. 22% for the Source subject in CND1), while in CND1, 78% of all personal pronouns were used to pick up the non-subject=Goal antecedent (vs. 7% for the Source object in CND2).

3 Conclusion

The experimental study revealed two main findings. First, grammatical role has a strong impact on the type of referring expression used. As many approaches to accessibility predict,

participants produced more null pronouns when referring to the subject of the previous sentence, than to other referents. The Romanian null pronouns are presumably the default option for resuming the subject antecedent. This result is in line with previous observations on Romanian and the predictions.

Second, the overt personal pronouns were more versatile, being used to pick up both the subject and non-subject referent, not being specialized for reference to non-subjects, as expected. So, null and overt Romanian pronouns do not exhibit a division of labour effect such that a subject bias for null pronouns corresponds to a non-subject bias for overt pronouns.

Third, this study brings favourable evidence for the observation that the use of a particular thematic role affects the choice of subsequent mention (Arnold 2001, Kaiser, Li & Holsinger 2011, Rosa & Arnold 2017). Participants produced more (null and overt) pronouns to refer to the previous Goal than to the previous Source. The Goal-bias was stronger for the non-subject, compared to the subject. In addition, the Goal-bias changed the observed preference of overt pronouns to pick up the subject referent, as for this condition more overt pronouns were used for the non-subject than for the subject referent. The results predict that Goals are more accessible than Sources, which contributes to the use of more reduced types of referring expressions. Thematic role effects are strongest in overt pronoun production in Romanian.

Overall, this is the first study to show that subjecthood alone does not suffice to account for pronoun production in Romanian. As in English, thematic roles may affect reference form production in addition to grammatical role. The present results furthermore supports a multi-dimensional approach, suggesting that different referential forms are constrained by different grammatical and semantic factors and that these factors might differ cross-linguistically (Ariel 1990, Kaiser & Trueswell 2008; von Heusinger & Chiriacescu 2009, Chiriacescu & von Heusinger 2011, Lindemann 2020).

Author Note

References

Ariel, Mira. 1990. *Accessing noun-phrase antecedents*. London: Routledge.

Arnold, Jennifer. 2001. *Reference form and discourse patterns*. PhD dissertation, Stanford University.

Carminati, Nella. 2002. *The processing of Italian subject pronouns*. PhD dissertation, University of Massachusetts Amherst.

Chiriacescu, Sofiana I. & Klaus von Heusinger. 2011 The discourse structuring potential of definite noun phrases in natural discourse. In Iris Hendrickx, António Branco, Sobha Lalitha Devi & Ruslan Mitkov (eds.), *Anaphora and reference resolution: Proceedings of the 8th discourse anaphora and anaphor resolution colloquium*, 151–162. Lisbon: Edições Colibri.

Chiriacescu, Sofiana. 2011. Factors contributing to the salience of referents. In Manuela Nevaci (ed.), *Studia linguistica et philologica*, 74–96. Bucuresti: Editura Universitatii.

Dobrovie-Sorin, Carmen. 1994. *The Syntax of Romanian*. Berlin: De Gruyter.

Fukumura, Kumiko & Roger van Gompel. 2010. Choosing anaphoric expression. Do people take into account likelihood of reference? *Journal of Memory and Language* 62. 52–66.

Givón, Talmy. 1981. On the development of the numeral 'one' as an indefinite marker. *Folia linguistica historica* 2. 35–53.

Grosz, Barbara, Aravind Joshi & Scott Weinstein. 1995. Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics* 21. 203–225.

Gundel, Jeanette, Nancy Hedberg & Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69(2). 274–307.

von Heusinger, Klaus & Sofiana Chiriacescu. 2009. Definite "bare" nouns and *pe*-marking in Romanian. In M. Teresa Espinal, Manuel Leonetti & Louise McNally (eds.), *Proceedings of the IV Nereus International Workshop "Definiteness and DP Structure in Romance Languages"*, 63–82. Konstanz: Universität Konstanz.

Kaiser, Elsi & John Trueswell. 2008. Interpreting pronouns and demonstrantives in Finnish: Evidence from a form-specific approach to reference resolution. *Language and Cognitive Processes* 23(5). 709–748.

Kaiser, Elsi, David Li & Edward Holsinger. 2011. Exploring the lexical and acoustic consequences of referential predictability. In Iris Hendricks, António Branco, Sobha Lalitha Devi & Ruslan Mitkov (eds.), *Anaphora Processing and Applications, Lecture Notes in Artificial Intelligence*, 171–183. Heidelberg: Springer.

Kehler, Andrew, Laura Kertz, Hannah Rohde & Jeffrey Elman. 2008. Coherence and coreference revisited. *Journal of semantics* 25. 1–44.

Lindemann, Sofiana. 2020. *Special indefinites in sentence and discourse*. Tübingen: Narr Franke Attempto.

Pagurschi, Florentina. 2010. *Interpreting Pronouns in Discourse Representation Theory*. Doctoral dissertation. Bucharest: University of Bucharest Press.

Rosa, Elise & Jennifer Arnold. 2017. Predictability affects production: Thematic roles can affect reference form selection. *Journal of Memory and Language* 94. 43–60.

Teodorescu, Otilia. 2016. On pronominal anaphora resolution in temporal adjuncts in chid Romanian. *Working Papers in Linguistics* XVIII. 37–48.

Zafiu, Rodica. 2008. Anafora. In Valeria Guţu Romalo (ed.), *Gramatica limbii române II. Enunţul*, 656–672. Bucharest: Editura Academiei Române.