

The Time Course of Pronoun Resolution in Natural Text Reading

Olga Seminck & Pascal Amsili

Laboratoire de Linguistique Formelle (Université Paris Diderot & CNRS)

`olga.seminck@cri-paris.org`, `pascal.amsili@linguist.univ-paris-diderot.fr`

In the study reported here, we model reading times relative to pronoun resolution. Instead of using hand-constructed experimental material, we study pronoun resolution in natural text reading. We present a series of models of the first pass reading time on the Dundee Corpus (Kennedy et al., 2003). This abstract presents our method, and some preliminary results. On the one hand, we show that the models confirm factors of influence described in a large psycholinguistic literature. On the other hand, we present how the models can be used to study new hypotheses about factors of influence on pronoun resolution.¹

Method

We use the English part of the Dundee Corpus (Kennedy et al., 2003), a corpus of ~50K words from newspaper articles that was read by 10 people whose eye movements were recorded. We also use an additional annotation layer: all the 1109 anaphoric pronouns of the corpus are annotated with their closest antecedent (Seminck and Amsili, 2018)². For each of these pronouns we modeled log-transformed first pass reading times, for six different regions of interest: the word before the pronoun, the pronoun itself and the four words following it. We did this for two reasons. First: pronouns are often not fixated, because they are short words. Nevertheless they are often read in parafoveal vision, when a fixation falls on an adjacent word. Second: the effect of pronoun resolution might be delayed, showing up only some time after the pronoun is read (Ehrlich and Rayner, 1983; Van Gompel and Majid, 2004). In sum, the effect of pronoun resolution is likely not to show only in fixations on the pronoun.

Each region corresponds to a mixed-effects model from the `lme4` R package (Bates et al., 2015). We modeled each of the 1109 pronouns and each of the 10 participants with random intercepts. Besides intercepts, our models contain two types of factors: control factors and pronoun factors. Pronoun factors are factors that deal with the anaphoric relation between the pronoun and its antecedent. These factors will be discussed in more details in the next section. Control factors, on the other hand, are low-level factors — for example word length — that are known to have an influence on reading time. In our study we controlled for forward and backward probability from an n-gram model (Frank and Bod, 2011). We also controlled for the length in characters of the reading region. We controlled for word frequency with the log-frequency both in the Dundee Corpus and in the British National Corpus of the word in the region and the word previous to the region. Finally, we also added the launch and the landing position of the first fixation on the region to the control factors. In the interest of space, these control factors will not be discussed further in this abstract.

¹Acknowledgments: Thanks to Yair Haendler and Barbara Hemforth for their advise. We also thank the three anonymous reviewers for their comments. This work was supported by the Labex EFL (ANR-10- LABX-0083).

²This resource is freely available at <http://www.llf.cnrs.fr/apadec>.

Factors

Our study has two goals. A first goal of our research is to seek confirmation for a large number of factors of influence described in the psycholinguistic literature. We studied the distance between the pronoun and its antecedent (e.g. Ehrlich and Rayner, 1983), the frequency of the antecedent (e.g. Pynte and Colonna, 2000; Van Gompel and Majid, 2004), *the subject preference* (e.g. Crawley et al., 1990), and *grammatical parallelism* (e.g. Smyth, 1994). A second goal of our research is to use our corpus as a means to discover new factors and to study less well known factors. We studied for example the difference between various gender and number features. We studied whether there was a difference in resolution time between masculine, feminine, neutral and plural pronouns. We found that the neutral gender slows down the reading, while the plural leads to a speed up. Furthermore, we investigated the hypothesis that a larger number of antecedent candidates in a text leads to more processing cost of anaphora (Seminck and Amsili, 2017). As a first approach to test this hypothesis, we measured the distance from the anaphoric pronoun to the beginning of a text. It is expected that, as the text unfolds, more characters are introduced, augmenting the number of potential antecedents. We found an effect of this factor: more distance leading to more reading time, a result in line with the hypothesis. Finally we also tested whether the length of the antecedent in words had an influence on the reading time of the pronoun, because according to the Accessibility Theory (Ariel, 1991), more salient discourse referents are marked in a more compact way. But we did not find significant results with this last factor. Our results are presented in somewhat more details in Table 1.

Results

Table 1 gives the model estimates for each region and each factor. The factors inspired by the literature confirm findings from previous experiments. We find that an antecedent that has a higher frequency in the corpus leads to less reading time (log freq antecedent Dundee, region 4, $p < 0.05$), that a pronoun in the subject position is read faster (syntactic role pro subj, region 2, $p < 0.05$) and that antecedents with syntactic roles other than subject and direct object slow down the reading time of the pronoun (syntactic role antecedent other, region 1, $p < 0.05$). The only factor from literature that behaves slightly differently is the distance between the pronoun and its antecedent. For this factor we first observe that the word before the pronoun is read faster when the distance is longer (distance antecedent & pronoun, region -1, $p < 0.05$) and that the reading is delayed later (region 3, $p < 0.05$). One way to interpret this pattern is the following: pronouns at a longer distance are resolved later. It could be that pronouns that are close to their antecedents are resolved immediately when they are fixated, but that when the distance is longer, the cognitive load of the resolution shows on later fixations.

As for new factors, we see that the neutral gender (*it*) leads to more reading time (region 0, $p < 0.05$). This effect can be explained by the fact that *it* can be anaphoric and pleonastic, but all the occurrences we studied were anaphoric. Hence, the increased reading time can come from the decision readers have to take whether the pronoun is anaphoric or not. On the contrary, we found that the plural pronoun (*they*) leads to a speed up (region 1, $p < 0.05$) compared to the other pronouns. We think that *they* might be less ambiguous than other pronouns. In future work we will test this hypothesis. For the distance from the pronoun to the beginning of the text we can see a significant positive effect (distance pronoun begin region 1, $p < 0.05$), suggesting that that more discourse referents increase the processing load of anaphora. However, this finding must be confirmed with future experiments. For the length of the antecedent no significant effect was found.

Table 1: Coefficient estimates for the models of first pass reading time for the six regions of interest (from one word before the pronoun to four words after, the pronoun being region 0). Control factors and intercepts are not reported.

Region	-1	0	1	2	3	4
Literature Factors						
distance antecedent & pronoun	-4.1E-03 *	2.2E-03	6.4E-04	1.1E-03	4.2E-03 *	2.4E-03
log freq antecedent Dundee	1.8E-03	2.3E-03	2.8E-03	2.7E-03	-5.5E-03	-8.5E-03 *
log freq antecedent BNC	-2.5E-04	1.4E-03	-4.8E-04	-7.2E-05	3.8E-03	5.6E-03
syntactic role pro subj	6.4E-03	8.1E-03	-7.1E-03	-2.0E-02 *	-7.5E-03	-2.7E-03
syntactic role pro other	9.3E-04	5.3E-03	4.2E-04	-3.3E-03	-1.2E-02	-9.5E-03
syntactic role antecedent subj	7.2E-04	3.7E-03	1.4E-02	4.8E-03	1.1E-02	1.1E-02
syntactic role antecedent other	1.6E-03	-1.3E-03	2.0E-02 *	6.7E-04	1.3E-02	-6.6E-03
parallel function	-5.6E-03	-1.3E-02	-2.6E-03	-1.2E-04	-2.8E-03	-1.1E-02
New Factors						
pronoun = they	1.3E-03	1.5E-02	-1.9E-02 *	-8.5E-03	-1.2E-02	7.4E-03
pronoun = it	-1.0E-02	2.5E-02 *	-3.3E-03	-4.9E-03	1.7E-03	1.2E-02
distance pronoun begin	-3.3E-03	-2.4E-04	4.7E-03 *	7.6E-04	-9.2E-05	-6.3E-04
length antecedent	1.0E-03	7.2E-04	3.6E-04	1.4E-03	-1.1E-03	-2.1E-04

Significance : *** for $p < 0.001$; ** for $p < 0.01$; * for $p < 0.05$ and . for $p < 0.1$.

Conclusion

Our study on pronoun resolution is — to our knowledge — the first using data from natural text reading. With our data and method, we were able to confirm the influence of factors known from a vast psycholinguistic literature on anaphora resolution, adding evidence for the robustness of these factors. Furthermore, our method allowed us to study new factors of influence. In the future, our models can be improved by adding (automatically) more annotation to the pronoun-antecedent pairs in our resource. For example, we could add an annotation of implicit causality to the verbs in the Dundee Corpus.

References

- Mira Ariel. The function of accessibility in a theory of grammar. *Journal of Pragmatics*, 16(5):443–463, 1991.
- Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48, 2015. doi: 10.18637/jss.v067.i01.
- Rosalind A. Crawley, Rosemary J. Stevenson, and David Kleinman. The use of heuristic strategies in the interpretation of pronouns. *Journal of Psycholinguistic Research*, 19(4):245–264, 1990.
- Kate Ehrlich and Keith Rayner. Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behavior*, 22(1): 75–87, 1983.
- Alan Kennedy, Robin Hill, and Joël Pynte. The Dundee Corpus. In *Proceedings of the 12th European conference on eye movement*, 2003.
- Joël Pynte and Saveria Colonna. Decoupling syntactic parsing from visual inspection: The case of relative clause attachment in french. *Reading as a perceptual process*, pages 529–547, 2000.
- Olga Seminck and Pascal Amsili. A computational model of human preferences for pronoun resolution. In *Proceedings of the Student Research Workshop of EACL*, pages 53–63, 2017.
- Olga Seminck and Pascal Amsili. A gold anaphora annotation layer on an eye movement corpus. In *Proceedings of the 11th International Conference on Language Resources and Evaluation*, 2018.
- Ron Smyth. Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research*, 23(3):197–229, 1994.
- Roger P. G. Van Gompel and Asifa Majid. Antecedent frequency effects during the processing of pronouns. *Cognition*, 90(3):255–264, 2004.