

EFFECT OF RELATIVE PRONOUN TYPE ON RELATIVE CLAUSE ATTACHMENT

Claire Delle Luche¹, Roger P. G. van Gompel², Frédérique Gayraud¹, & Bruno Martinie¹

1 – Laboratoire Dynamique du Langage, Lyon, France.

2 – University of Dundee, United Kingdom.

ABSTRACT

Accessibility hierarchies (Ariel, 1990, 2001; Givón, 1992; Gundel, Hedberg, & Zacharski, 1993) assume that the form of anaphoric expressions signals the relative saliency of the antecedent. We argue that the form of relative pronouns in relative clauses has a similar function and therefore influences attachment preferences. We conducted two questionnaire experiments in which we investigated whether attachment preferences for ambiguous relative clauses are affected by the type of relative pronoun that is used. Experiment 1 showed a difference in attachment preference between *qui* and *lequel*, indicating that the form of the relative pronoun affects attachment preferences. Experiment 2 demonstrated that the difference observed in Experiment 1 is not due to differences in informativity between *qui* and *lequel*, suggesting that instead, it is due to a difference in markedness (*qui* is more frequent and shorter).

INTRODUCTION

Many functional linguistic theories assume that the form of anaphoric expressions signals how accessible their antecedent is (e.g., Ariel, 1990, 2001; Givón, 1992; Gundel, Hedberg, & Zacharski, 1993). These accessibility theories claim that anaphoric expressions can be ranked according to the accessibility of the antecedent that they tend to refer to. For example, pronouns signal that the antecedent is highly accessible, whereas noun phrases and names signal that the antecedent is relatively inaccessible. Hence, pronouns are ranked higher on the accessibility hierarchy than noun phrases and names. More generally, short anaphoric expressions that provide little semantic and syntactic information tend to be high on the accessibility hierarchy, whereas longer expressions that contain more information about their antecedent tend to be low on the accessibility hierarchy.

Most evidence for accessibility hierarchies comes from corpus studies, which show that in conversations and texts, people tend to use anaphoric expressions that are high on the accessibility hierarchy when referring to very accessible antecedents, but expressions low on the hierarchy when referring to antecedents that are inaccessible (Givón, 1992; Gundel et al., 1993). For example, Gundel et al. (1993) analysed the distribution of different anaphoric expressions and observed that pronouns were more frequently used when the antecedent was in focus and therefore highly accessible, whereas definite noun phrases were mostly used to refer to uniquely identifiable antecedents that were not in focus.

Accessibility hierarchies also receive support from reading studies. In a series of experiments, Gordon and colleagues (Gordon, Grosz, & Gilliom, 1993; Gordon & Chan, 1995) have shown that names are harder to process than pronouns when they refer back to

an antecedent name that is the subject of the preceding sentence. This effect has been dubbed the *repeated name penalty*. The repeated name penalty is affected by the syntactic role of the antecedent noun phrase: When the antecedent name is a direct object, the penalty is much reduced. The repeated name penalty effect has been taken to support *centering theory*, a computational theory accounting for coherence in texts (Grosz, Joshi, & Weinstein, 1983, 1995). More generally, it provides support for the idea that the preference for a particular anaphoric expression is affected by the saliency of the antecedent. When the antecedent has the syntactically highly salient role of subject, an anaphor that is high on the accessibility hierarchy such as a pronoun is easier to process than a name, which is low on the hierarchy. But when the antecedent is a direct object and therefore less salient, the processing advantage for pronouns relative to names disappears.

An interesting question is whether relative pronouns can also be ranked on the accessibility hierarchy. As suggested by the term relative *pronoun*, they can be considered to be a type of anaphor. This idea is consistent with Hemforth, Konieczny, and Scheepers (2000), who argued that the processing of relative clauses (RCs henceforth) involves both syntactic attachment of the RC into the preceding tree structure and anaphoric binding of the relative pronoun. If this is true and relative pronouns indeed behave similarly to personal pronouns, we expect that relative pronouns also signal how accessible their antecedent is. This should have an effect on how people process ambiguous RCs such as (1), which have been investigated in much psycholinguistic research.

(1) *The journalist interviewed the daughter of the colonel who had had the accident.*

A number of studies have shown that in English, the RC *who had had the accident* in (1) is preferentially interpreted as modifying the second noun phrase (NP2 henceforth) *the colonel* rather than *the daughter* (NP1) (e.g., Carreiras & Clifton, 1999; Cuetos & Mitchell, 1988), while other studies suggest that there is no strong preference for either analysis (Carreiras & Clifton, 1993; Traxler, Pickering, & Clifton, 1998). This is consistent with the late closure principle (Frazier, 1979, 1987), which claims that the ambiguous RC should be attached as low as possible into the preceding tree structure, and is also compatible with a recency principle, which predicts that it should be attached to the most recent phrase (e.g., Gibson, Pearlmutter, Canseco-Gonzalez, & Hickok, 1996; Stevenson, 1994). By contrast, in languages such as French, Spanish, German, and Dutch, there is a preference for attachment to NP1. There has been much debate about the reasons behind the NP1 attachment preference in these languages and the reasons behind the cross-linguistic differences. One possibility, suggested by Mitchell, Cuetos, Corley, and Brysbaert (1995) is that NP2 attachment is more frequent in English, whereas NP1 attachment occurs more frequently in languages such as Spanish and Dutch. By contrast, Frazier and Clifton (1996) argued that pragmatic principles affect RC attachment differently in different languages, while Gibson et al. (1996) argued that recency, which favours attachment to the most recent noun phrase (NP2 in 1) and predicate proximity, which favours attachment to the head of the predicate (NP1) have different weights in different languages. Finally, Hemforth et al. (2000) claimed that in languages like German and Spanish, relative pronouns are processed like personal pronouns and are therefore preferentially interpreted as coreferent with the most salient NP, that is, NP1. By contrast, in English, relative pronouns are often omitted or generalised complementisers (*that*), so the parser relies more on syntactic processing strategies favouring NP2 attachment.

The current study did not aim to distinguish between the different theories of RC attachment, but instead, it aimed to investigate whether the form of the relative pronoun affects RC attachment preferences. Until now, none of the theories has considered the

possibility that the type of relative pronoun might affect RC processing, and no studies have addressed this question. However, if accessibility hierarchies generalise to relative pronouns, we expect that attachment preferences for ambiguous RCs should be affected by the type of relative pronoun: RCs with relative pronouns that signal a highly accessible antecedent should preferentially attach to the most accessible NP, whereas this preference should be less strong for RCs with relative pronouns that signal a less accessible antecedent.

In order to investigate the influence of relative pronouns on RC attachment preferences, we will compare two types of relative pronouns in French. It has been demonstrated (Zagar, Pynte, & Rativeau, 1997) that in sentences like (2) containing a *qui* RC, readers of French prefer to attach the RC *qui semblait plus confiant* to *l'avocat* (NP1 attachment) rather than to *la chanteuse* (NP2 attachment).

(2) *Un journaliste aborda l'avocat de la chanteuse qui semblait plus confiant.* (A journalist approached the barrister_{MASC} of the singer_{FEM} who seemed more confident_{MASC}.)

However, French also has a different type of relative pronoun, namely *lequel* or *laquelle*. *Lequel* and *laquelle* may be lower on the accessibility hierarchy than *qui* because *lequel* and *laquelle* are more marked, that is, they are phonologically longer and less frequent than *qui*. Furthermore, *lequel* and *laquelle* are marked for gender and number, so they are also more informative than *qui*, which does not have gender and number marking. Both factors may affect the position of the relative pronoun on the accessibility hierarchy (Ariel, 1990, 2001). Therefore, RCs with *lequel* or *laquelle* should be attached to less accessible NPs. Therefore, RCs with *lequel* or *laquelle* should be attached to less salient NPs. Assuming that NP1 is most salient, this predicts that in a sentence like (3), readers should make a local attachment to NP2 more frequently, resulting in a less strong NP1 attachment preference than in *qui* RCs.

(3) *Un journaliste aborda l'avocate de la chanteuse, laquelle semblait plus confiante.* (A journalist approached the barrister_{FEM} of the singer_{FEM} who seemed more confident_{FEM}.)

EXPERIMENT 1: QUI VS. LEQUEL/LAQUELLE

The purpose of Experiment 1 was to test whether the type of relative pronoun affects RC attachment in French. We compared *qui* (1) and *lequel/laquelle* (2) in non-restrictive RCs following a 'NP1 of the NP2' structure. The accessibility account predicts that *qui* RCs should preferentially attach to the most accessible NP (presumably NP1, e.g., Zagar, Pynte, & Rativeau, 1997), but this preference should be weaker for *lequel/laquelle* RCs. This is because *qui* is less marked and less informative than *lequel/laquelle*, so *qui* should refer to more accessible antecedents (Ariel, 1990).

METHOD

Participants Fifty-six undergraduates at Lyon 2 Lumière University participated in the experiment as part of their courses. All were native speakers of French.

Materials and procedure Thirty-two sets of experimental sentences were constructed in two versions, one with *qui* RCs (4), and one with the relative pronoun *lequel* or *laquelle* (5).

(4) *Je connais le père du maçon, qui est amusant.* (I know the father_{NP1} of the mason_{NP2} who is funny)

(5) *Je connais le père du maçon, lequel est amusant.* (identical meaning)

Each experimental trial consisted of a sentence such as (3) and (4) followed by two statements, one consistent with the NP1 attachment interpretation (*The father is funny*) and one consistent with NP2 attachment (*The mason is funny*). Participants were instructed to tick the option that was ‘most correct’. NP1 and NP2 were matched for gender, length and number of syllables.

Prior to the experiment, we conducted a pretest to ensure that the RC was not biased toward NP1 or NP2. Twenty participants, none of whom took part in Experiment 1, were asked to rate on a seven-point scale the plausibility of statements consistent with NP1 attachment and NP2 attachment. For the 32 sentences that we selected, there was no overall preference for either NP1 (5.83) or NP2 (5.95) attachment.

Two lists were constructed using a between subjects design: one list presented *qui* RCs, the other *lequel/laquelle* RCs. The experiment lasted about 20 minutes. The order of statements was counterbalanced for the two lists. The questionnaires were run in large groups of participants.

RESULTS AND DISCUSSION

We conducted two ANOVAs on the percentages of NP1 attachment responses, one with subjects ($F1$) and one with items ($F2$) as the random variable. The ANOVAs contained condition (*qui* vs. *lequel/laquelle*) as a between subjects and within items variable. Figure 1 presents the mean percentage of NP1 attachment by condition. The results showed that for *qui* RCs, participants strongly preferred NP1 (87.2% of trials), but this preference was much weaker (70.6%) with *lequel/laquelle* RCs. The percentage of NP1 attachments differed from chance in both the *qui* RCs ($F1(1,27) = 17.70$; $p < .01$; $F2(1,31) = 26.51$; $p < .01$) and the *lequel/laquelle* RCs ($F1(1,27) = 4.40$; $p < .01$; $F2(1,31) = 12.64$; $p < .01$). Most important, the difference between conditions was significant both by subjects ($F1(1,54) = 9.01$; $p < .01$) and items ($F2(1,31) = 4.73$; $p = .04$).

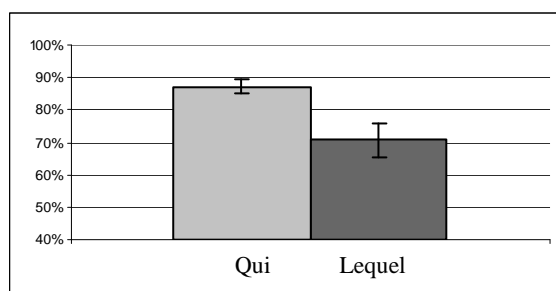


Figure 1: Mean NP1 attachments (%)

EXPERIMENT 2: A QUI VS. AUQUEL

Qui may be higher on the accessibility hierarchy than *lequel/laquelle* because (1) *qui* is less marked than *lequel/laquelle*, that is, *qui* is phonologically shorter and more frequent than *lequel/laquelle* and (2) because *qui* is less informative than *lequel/laquelle*, that is, *lequel/laquelle* are marked for gender and number, whereas *qui* is not. Ariel (1990, 2001) argued that both factors may affect an anaphor's position on the accessibility hierarchy, so both factors may affect RC attachment preferences.

The goal of Experiment 2 was to test whether the difference in attachment preferences we found in Experiment 1 were due to a difference in markedness between *qui* and *lequel/laquelle* or due to a difference in informativity between *qui* and *lequel/laquelle*. In the current experiment, we controlled the relative pronouns for phonological length by comparing dative RCs containing either *à qui* or *auquel*. These relative pronouns are also more similar in frequency than *qui* and *lequel/laquelle* in Experiment 1. However, they differ in informativity: *auquel* is marked for gender (masculine) and number (singular), whereas *à qui* is not. Hence, if *à qui* and *auquel* RCs have different attachment preferences, this must be due to the difference in informativity. By contrast, if they have the same attachment preferences, this suggests that the difference in attachment preference in Experiment 1 was due to a difference in markedness (length and frequency differences between *qui* and *lequel/laquelle*).

METHOD

Participants Twenty-four participants from the same population took part in this experiment. None had participated in Experiment 1 or the pretests.

Materials and procedure The method and design were similar to those in Experiment 1. On the basis of norms from a plausibility pretest (20 participants, same procedure as the pretest in Experiment 1), we selected 32 sentences that did not differ in plausibility between conditions (NP1, 4.87; NP2, 4.91). Sentences contained either *à qui* (6) or *auquel* (7).

(6) *Je connais le collègue de l'étudiant, à qui la bibliothécaire apporte le livre. (I know the colleague of the student, to whom the librarian is giving the book)*

(7) *Je connais le collègue de l'étudiant, auquel la bibliothécaire apporte le livre. (identical meaning)*

We used the same fillers as in Experiment 1. Two lists were constructed using a between subjects design.

RESULTS AND DISCUSSION

As in Experiment 1, we conducted analyses by subjects and items on the percentages of NP1 attachment responses with the variable condition as a between subjects and within items variable. Figure 2 presents the mean attachment preference by condition. The mean percentages of NP1 attachment were high (Figure 2) both with *à qui* (87.6%) and *auquel* (85.3%). The percentage of NP1 attachments differed from chance for *qui* RCs ($F(1, 11) = 10.05$; $p < .01$; $F(1,31) = 10.52$; $p < .01$) as well as *auquel* RCs ($F(1,11) = 4.56$; $p < .01$; $F(1,31) = 22.32$; $p < .01$). No significant difference between the two conditions was found ($F_s < 1$): Participants chose NP1 equally often with *auquel* RCs as with *à qui* RCs.

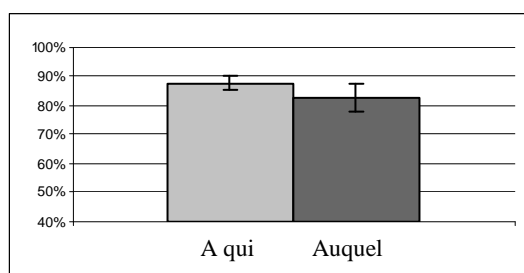


Figure 2: Mean NP1 attachments (%)

The results of Experiment 2 demonstrate that the difference in informativity between *à qui* and *auquel* did not contribute to differences in attachment preferences. This suggests that the difference observed in Experiment 1 must be due to a difference in markedness. The position of the relative pronoun on the accessibility hierarchy is thus affected by markedness rather than informativity.

GENERAL DISCUSSION

Two experiments investigated whether and how RC attachment preferences are affected by the form of the relative pronoun. Experiment 1 showed a strong NP1 attachment preference when the relative pronoun was *qui*, but this preference was much reduced when it was *lequel* or *laquelle*. This result provides support for the idea that relative pronouns can be ranked on the accessibility hierarchy. Relative pronouns such as *qui* signal highly accessible antecedents, and are therefore preferentially interpreted as referring to NP1, whereas relative pronouns such as *lequel* and *laquelle* signal less accessible antecedents, so the preference for NP1 attachment is less strong.

An important question is what factors determine a relative pronoun's position on the accessibility hierarchy. *Lequel* and *laquelle* are more marked than *qui*, because they are longer and less frequent than *qui*. But in addition, they are also more informative than *qui*, because they contain gender and number marking. In Experiment 2, we controlled for markedness of the relative pronoun by contrasting *à qui* and *auquel* (which are similar in length and frequency), while manipulating informativity (gender and number marking). The experiment demonstrated that *auquel* and *à qui* had similarly strong NP1 attachment preferences despite the fact that they differ in informativity. Therefore, in Experiment 1, it was markedness rather than informativity that contributed to the difference in attachment preferences. Hence, our experiments suggest that markedness of the relative pronoun (resulting from a low frequency of the relative pronoun and its length) affects the position of the relative pronoun on the accessibility hierarchy, whereas informativity does not.

Our results have implications for both theories of anaphors and sentence processing theories. Relative pronouns appear to have properties that are similar to those of other anaphors such as personal pronouns. Similar to other anaphors, relative pronouns refer back to an earlier introduced entity in the discourse. And like other anaphors, their form influences how they are preferentially interpreted. When a relative pronoun is unmarked, that is, it is short and frequent, it signals a highly accessible antecedent. But when it has a marked form, it signals a less accessible antecedent. Therefore, unmarked relative pronouns are preferentially interpreted as referring to the most salient antecedent NP, whereas this preference is less strong for marked relative pronouns. Of course, this is not to say that relative pronouns are similar to personal pronouns in all respects. Clearly, syntactic constraints on relative pronouns are different from constraints on pronouns. For example, unlike personal pronouns, relative pronouns must be in the same sentence as their antecedent, and unlike personal pronouns, they cannot precede their antecedent in English. However, this is not surprising: It is well-known that syntactic constraints on personal pronouns and noun phrase anaphors are also different (e.g., Chomsky, 1981, Reinhart, 1983). What we would like to argue is that there are good arguments to believe that relative pronouns belong to the class of anaphoric expressions and that their processing is affected by similar factors.

Our results have important implications for sentence processing theories too, because they showed that parsing preferences are affected by the form of the relative pronoun. This is difficult to reconcile with many sentence processing theories, because they do not assign a role to the form of the relative pronoun. For example, the garden-path theory (Frazier, 1979, 1987) predicts a preference for NP2 attachment due to the application of the late closure

strategy. This does not explain why French exhibits a NP1 attachment preference, and furthermore, given that late closure is a purely syntactic strategy, it does not explain why attachment preferences are affected by the type of relative pronoun. Gibson et al. (1996) argued that NP1 or NP2 attachment is preferred depending on whether predicate proximity favouring NP1 attachment or recency favouring NP2 attachment is the strongest parsing constraint. This would explain the current results if it is assumed that recency is a stronger constraint for *lequel/laquelle* than for *qui*. However, there does not seem to be a principled reason why this should be the case. Frequency-based accounts (e.g., Desmet, De Baecke, Drieghe, Brysbaert, & Vonk, in press; Mitchell et al., 1995) fare slightly better, because it seems likely that *qui* is used more often to refer to NP1 than *lequel/laquelle*. But this raises the question: Why do these production preferences occur? Accessibility theories have an answer to this. Language producers signal that the antecedent of a relative pronoun is highly accessible by using an unmarked relative pronoun that is high on the accessibility hierarchy, whereas they signal that the antecedent is relatively inaccessible by using a marked relative pronoun that is low on the hierarchy. The results from our two experiments show that comprehenders use these accessibility cues, and therefore, the form of the relative pronoun affects the comprehension of RCs.

REFERENCES

- Almor, A. (1999). Noun-phrase anaphora and focus: The informational load hypothesis. *Psychological Review*, 106, 748-765.
- Ariel, M. (1990). *Accessing noun-phrase antecedents*. London: Croom Helm.
- Ariel, M. (2001). Accessibility theory: An overview. In T. J. M. Sanders (Ed.), *Text representation: Linguistic and psycholinguistic aspects*. Philadelphia, PA, USA: John Benjamins.
- Carreiras, M., & Clifton, C. (1993). Relative clause interpretation preferences in Spanish and English. *Language and Speech*, 36, 353-372.
- Cuetos, F., & Mitchell, D. C. (1988). Cross-linguistic differences in parsing: Restrictions on the use of the Late Closure strategy in Spanish. *Cognition*, 30, 73-105.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.
- Desmet, T., Brysbaert, M., & De Baecke, C. (2002). The correspondence between sentence production and corpus frequencies in modifier attachment. *Quarterly Journal of Experimental Psychology*, 55A, 879-896.
- Frazier, L. (1979). *On comprehending sentences: Syntactic parsing strategies*. Ph.D. Dissertation. Indiana University Linguistics Club. University of Connecticut.
- Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 559-586). Hillsdale, NJ: Erlbaum.
- Frazier, L., & Clifton, C., Jr. (1996). *Construal*. Cambridge, MA: MIT Press.
- Gibson, E., Pearlmutter, N., Canseco Gonzalez, E., & Hickok, G. (1996). Recency preference in the human sentence processing mechanism. *Cognition*, 59, 23-59.
- Givón, T. (1992). The grammar of referential coherence as mental processing instructions. *Linguistics*, 30, 5-55.
- Gordon, P. C., & Chan, D. (1995). Pronouns, passives, and discourse coherence. *Journal of Memory and Language*, 34, 216-231.
- Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 69, 274-307.

Gordon, P. C., Grosz, B., & Gilliom, L. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, 3, 311-347.

Gosz, B. J., Joshi, A. K., & Weinstein, S. (1983). Providing a unified account of definite noun phrases in discourse. *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics*, Cambridge, MA.

Gosz, B. J., Joshi, A. K., & Weinstein, S. (1995). Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics*, 21, 203-225.

Hemforth, B., Konieczny, L., & Scheepers, C. (2000). Syntactic attachment and anaphor resolution: The two sides of relative clause attachment. In M. W. Crocker, M. J. Pickering, & C. E. Clifton, Jr. (Eds.), *Architectures and mechanisms for language processing* (pp. 259-282). Cambridge, UK: Cambridge University Press.

Mitchell, D.C., Cuetos, F., Corley, M.M.B., & Brysbaert, M. (1995). Exposure-based models of human parsing: Evidence for the use of coarse-grained (nonlexical) statistical records. *Journal of Psycholinguistic Research*, 24, 469-488.

Reinhart, T. (1983). *Anaphora and Semantic Interpretation*. Chicago: University of Chicago Press.

Stevenson, S. (1994). Competition and recency in a hybrid network model of syntactic disambiguation. *Journal of Psycholinguistic Research*, 23, 295-322.

Traxler, M.J., Pickering, M.J., & Clifton, C. (1998). Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language*, 39, 558-592.

Zagar, D., Pynte, J., & Rativeau, S. (1997). Evidence for early-closure attachment on first-pass reading times in French. *The Quarterly Journal of Experimental Psychology*, 50A, 421-438.