

Clausal interference during reference production

Roger P. G. van Gompel (r.p.g.vangompel@dundee.ac.uk)

School of Psychology, University of Dundee
Dundee, DD1 4HN, United Kingdom

Kumiko Fukumura (kumiko.fukumura@strath.ac.uk)

School of Psychological Science and Health, University of Strathclyde
Glasgow, G1 1QE, United Kingdom

Abstract

Four sentence completion experiments investigated how language producers refer back to human characters when they are mentioned in different clauses in the preceding sentence. Participants produced fewer pronouns and more names when the characters were mentioned in the same type of clause (both main or both subordinate) than when they were mentioned in different clause types. This suggests that clauses with the same status interfere, thereby making the discourse entities less accessible. Participants also used more pronouns and fewer names when the antecedent was in a main than a subordinate clause and when it was the most recent rather than a more distant clause, indicating that main and recent clauses are most accessible.

Keywords: Language production; Reference; Pronouns; Names; Interference; Competition; Clause status; Main and subordinate clause; Recency.

Introduction

A general assumption in both theoretical linguistics and psycholinguistics is that language producers' choice of anaphoric expressions is strongly affected by the salience or accessibility of the referent (e.g., Ariel, 1990; Arnold, 2008; Fukumura & Van Gompel, 2010; Givón, 1983; Grosz, Joshi, & Weinstein, 1995; Gundel, Hedberg, & Zacharski, 1993). If the referent is highly salient or accessible, language producers favour anaphoric expressions such as pronouns that are short and provide little information about the referent, whereas if the referent is less accessible, they tend to use more informative expressions such as full names or definite noun phrases. Various saliency factors have been shown to influence the choice of anaphor. For example, speakers tend to produce more pronouns and fewer names when the referent is the subject than the object in the preceding sentence (Fletcher, 1984; Fukumura & Van Gompel, 2010; Stevenson, Kleinman, & Crawley, 1994), when the referent was more recently mentioned in the preceding discourse (Ariel, 1990; Givón, 1983), when the referent is animate rather than inanimate (Fukumura & Van Gompel, 2011) and when there is no competitor in the linguistic or visual context than when there is (Arnold & Griffin, 2007; Fukumura, Van Gompel, & Pickering, 2010).

One other factor that has frequently been argued to affect salience is clausal status, that is, the distinction between main and subordinate clauses. It is often assumed that main clauses provide the main assertion or foregrounded information in the sentence, whereas subordinate clauses

contain presupposed, backgrounded information (Levinson, 1983; Matthiessen & Thompson, 1988; Miltsakaki, 2002). This is supported by experimental evidence. In a study by Morrow (1985), participants tended to choose a character in the main rather than subordinate clause as the antecedent of an ambiguous pronoun. In line with this, Cooreman and Sanford (1996) and Price (2008) found that reading times were faster for sentences with pronouns that referred to a character in the main than the subordinate clause. Furthermore, several studies have found that participants referred more often to characters in the main than the subordinate clause (Cooreman & Sanford, 1996; Harris & Bates, 2002; Price, 2008). However, none of these studies examined whether clausal status also affects the choice of anaphoric expression.

The current study addresses the question of how clausal status influences anaphor choice. Assuming that information in the main clause is more salient than in the subordinate clause, language producers should produce more reduced expressions and fewer explicit expressions for characters in the main than the subordinate clause. However, clausal status may affect anaphor choice in other ways too. Recent research suggests that the choice of anaphor is affected by similarity-based interference between the referent and competitor entities. There is much evidence that language producers tend to avoid pronouns when the referent has the same gender as another character (the competitor) in the context (Arnold & Griffin, 2007; Clancy, 1980; Fletcher, 1984; Fukumura et al., 2010). Fukumura, Hyönä, and Scholfield (in press) argued that this is not purely due to ambiguity avoidance. They showed that this effect occurred even in Finnish, where pronouns are not marked for gender, so pronouns are ambiguous regardless of whether the competitor has the same or a different gender. They concluded that a competitor that is semantically similar to the referent due to its gender competes with the referent, reducing the referent's activation. This, in turn, results in more explicit anaphors. Other studies provide further evidence for similarity-based interference. Fukumura and Van Gompel (2011) showed that participants produced more explicit referring expressions (more definite noun phrases, fewer pronouns) when both the referent and a competitor were animate or both were inanimate than when they differed in animacy. Fukumura, Van Gompel, Harley, & Pickering (2011) found that speakers produced more explicit referring expressions when they referred to a person

who was in the same situation as a competitor (e.g., both on a horse) than when this person was in a different situation.

All these studies investigated interference due to semantic features of the referent and competitor, or semantic properties of the context. However, other types of similarity may also cause interference, including similarity between the clauses in which the referent and competitor are mentioned. For example, in (1a), both *Rachel* and *Douglas* are in a main clause, whereas in (1b), *Rachel* is in the main clause and *Douglas* is in a subordinate clause.

1a. Main-main clause:

Rachel explored the cave and Douglas was lying on the beach.

1b. Main-sub clause:

Rachel explored the cave while Douglas was lying on the beach.

If clausal interference occurs, then information in the first clause should be less accessible when the second clause is also a main clause (1a) than when it is not (1b). Therefore, when language producers refer to *Rachel*, they should use fewer pronouns and more repeated names in (1a) than (1b). Clausal interference also predicts stronger interference from *Rachel* for *Douglas* in (1a) than (1b), making *Douglas* less accessible in (1a) than (1b). However, *Douglas* in (1a) is in a main clause whereas it is in a subordinate clause in (1b), and this should make it *more accessible* in (1a) than in (1b). Because the clausal interference and clause status effects go in opposite directions, they may cancel each other out, resulting in no differences for reference to *Douglas*.

An alternative account of how clausal similarity may affect anaphor choice is the relative saliency account (cf. Fukumura & Van Gompel, 2010). According to this account, the saliency of the referent is not only determined by factors affecting its own saliency, but also by factors influencing the saliency of competitors: The higher the saliency of the competitors, the lower the relative saliency of the referent. This also predicts that *Rachel* should be more accessible in (1a) than (1b). Because *Douglas* is in the main clause in (1a), it should be more salient than in (1b), where it is in the subordinate clause. Therefore, the relative saliency of *Rachel* should be lower in (1a) than (1b). This should result in fewer pronouns and more names for *Rachel* in (1a) than (1b). Although this is the same prediction as that of the clausal interference account, the relative saliency account makes a different prediction for *Douglas*. Because *Rachel* is in the main clause in both cases, it should not affect the relative saliency of *Douglas* in (1a/b), so language producers should use more pronouns to refer to *Douglas* in (1a), where he is mentioned in the main clause, than (1b) when he is mentioned in the subordinate clause.

Experiment 1

We tested the predictions of the clausal interference and relative saliency accounts using sentences such as (1). Forty

participants were asked to provide a written continuation to the sentences, starting with a new sentence. Forty experimental materials were used, randomly interspersed between 86 fillers with a variety of structures. In addition to conditions (1a) and (1b), we had two additional conditions in which the words in the first and second clause were reversed to counterbalance for the specific words used in the two clauses. The conditions were manipulated as within-subject and –item variables.

Figure 1 shows the percentages of pronouns and repeated names for references to name 1 (*Rachel* in (1)) and name 2 (*Douglas*) in the main-main (1a) and main-subordinate (1b) conditions. Continuations in which the first word did not refer to either name were excluded.

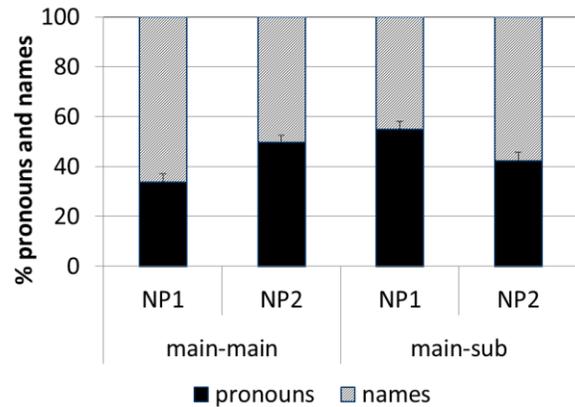


Figure 1: Results from Experiment 1.

By-subject and –item analyses of variance were conducted on arcsine transformed proportions of pronoun responses. We observed no main effect of recency (name 1 vs. 2) or clause structure (main-main vs. main-sub), but crucially, there was an interaction between these two variables ($F(1, 36) = 17.6, p < .001$; $F(1, 35) = 4.2, p = .048$). When referring to name 1, participants produced fewer pronouns in the main-main than main-sub condition ($F(1) = 20.9, p < .001$; $F(2) = 6.3, p = .017$), as predicted by both the clausal interference and the relative saliency account. In contrast, there was no effect of clause structure when referring to name 2 ($F(1) = 1.3, p = .26$; $F(2) < 1$). This is consistent with the clausal interference account, but does not support the relative saliency account, because the latter predicts that more pronouns should be used for name 2 in the main-main than main-subordinate condition.

Experiment 2

To further test the clausal interference and relative saliency accounts, we compared the main-main condition (2a) with a subordinate-main clause condition (2b).

2a. Main-main clause:

Rachel explored the cave and Douglas was lying on the beach.

2b. Sub-main clause:

While Rachel explored the cave Douglas was lying on the beach.

According to the clausal interference account, interference for *Douglas* should be larger in (2a) than (2b). In both conditions, *Douglas* is in a main clause, but it should get stronger interference from *Rachel*, when *Rachel* is also in a main clause (2a) than when *Rachel* is in a subordinate clause (2b). This should result in fewer pronouns and more names for *Douglas* in (2a) than (2b). Clausal interference for *Rachel* should also be larger in (2a) than (2b). However, *Rachel* is in the main clause in (2a) and in the subordinate clause in (2b), so the clause status effect is in the opposite direction from the clausal interference effect. The results from Experiment 1 suggest that they may cancel each other out.

The relative saliency account also predicts fewer pronouns for *Douglas* in (2a) than (2b). This is because the competitor *Rachel* is more salient in (2a), where it is in the main clause, than in (2b), where it is in the subordinate clause. As a result, the relative saliency of *Douglas* should be lower in (2a) than (2b). In contrast, the relative saliency of *Rachel* is higher in (2a) than (2b) because while competition from *Douglas* is the same in both conditions, *Rachel* is in the main clause in (2a) but in the subordinate clause in (2b). Therefore, participants should produce more pronouns for *Rachel* in (2a) than (2b).

Twenty-four participants took part. The method, design and analyses of the experiment were the same as before. Figure 2 shows the results by condition.

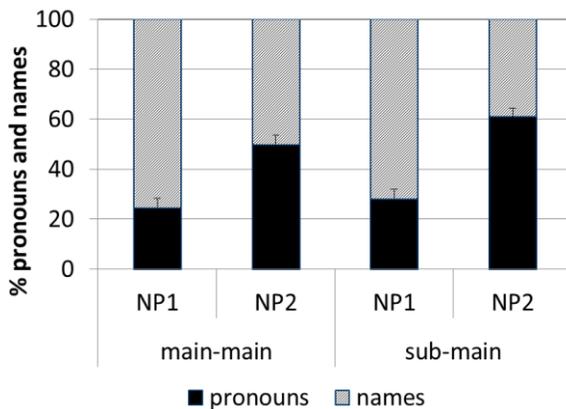


Figure 2: Results from Experiment 2.

We observed a main effect of recency, with more pronouns for the name in the more recent clause: $F_1 = 15.6, p < .001$; $F_2 = 83.7, p < .001$. When referring to name 2, participants produced fewer pronouns in the main-main than main-sub condition: $F_1 = 4.7, p = .043$; $F_2 = 1.1, p = .31$. When referring to name 2, there was no effect of clause structure: $F_s < 1$. These results are consistent with Experiment 1 and support the clausal interference account. The relative saliency account does not explain the results,

because it predicted more pronouns for name 1 (*Rachel*) in the main-main than sub-main condition.

Experiment 3

The clausal interference account explains the absence of an effect for name 2 in Experiment 1 and name 1 in Experiment 2 by assuming that the effects of clausal interference and clause status go in opposite directions and cancel each other out. In Experiment 3, we directly compared reference to main and subordinate clauses to provide evidence for an effect of clause status on the choice of anaphor. Although previous studies suggest that information in the main clause is more accessible than in the subordinate clause (Cooreman & Sanford, 1996; Harris & Bates, 2002; Price, 2008), we are not aware of research that shows how this affects anaphor choice. We used conditions with either a main-subordinate (3a) or subordinate-main clause order (3b).

3a. Main-sub order:

Rachel discovered the cave while Douglas was lying on the beach.

3b. Sub-main order:

While Rachel discovered the cave Douglas was lying on the beach.

Twenty-four participants took part. The method, design and analyses were the same as in the previous experiments. Figure 3 shows the results by condition.

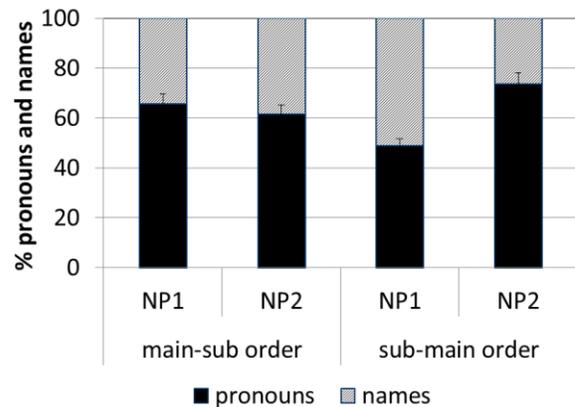


Figure 3: Results from Experiment 3.

We observed a main effect of clause status: Participants used more pronouns when they referred to the name in the main than the subordinate clause: $F_1 = 8.1, p = .010$; $F_2 = 10.6, p = .003$. There was also an effect of recency with more pronouns for the second than first clause: $F_1 = 11.9, p = .003$; $F_2(1, 31) = 3.4, p = .076$. Finally, there was an interaction between the two variables: In the sub-main order, there were more pronouns for the second than the first name ($F_1 = 16.5, p < .001$; $F_2 = 11.3, p = .002$), whereas there was no effect in the main-sub order ($F_s < 1$). Together, the results suggest that both clause status and

recency affect the choice of anaphor: Both variables make name 2 more accessible than name 1 in the sub-main order, whereas the effects cancel each other out in the main-sub order.

Experiment 4

Experiment 4 investigated whether clausal interference is stronger when the two clauses are syntactically closely linked. When two clauses are linked by a coordinating connective (4a), their mental representation in the syntactic and discourse structure may be closer than when there is no connective (4b). As a result, interference between the two referents *Rachel* and *Douglas* may be larger in (4a) than (4b). This should result in a main effect of clause linkage: The mean proportion of pronouns for the two referents should be lower in (4a) than (4b).

4a. Two coordinated clauses, strong link:

Rachel discovered the cave and Douglas was lying on the beach.

4b. Two sentences, no link:

Rachel discovered the cave. Douglas was lying on the beach.

In addition, we should get a recency effect, as in Experiment 3. This recency effect may be stronger when the two clauses are not linked by a connective (4b) and are therefore in separate processing units than when there is a connective (4a). This predicts an interaction between presence/absence of the connective and reference to the first or second name: More pronouns for name 2 in (4b) than (4a), but fewer pronouns for name 1 in (4b) than (4a).

Thirty-two participants took part. The method, design and analyses were the same as before. Figure 4 shows the results by condition.

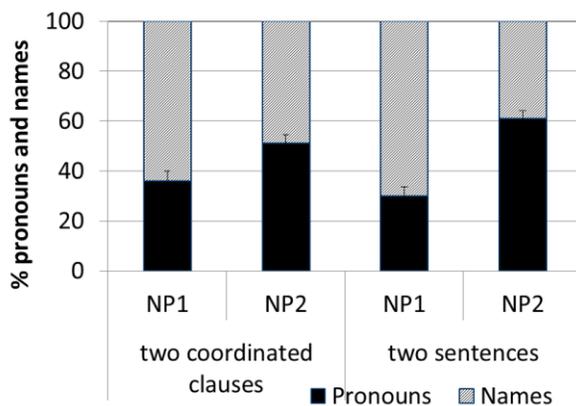


Figure 4: Results from Experiment 4.

The analyses showed a main effect of recency with more pronouns and fewer names for name 2 than 1: $F_1 = 20.9$, $p < .001$; $F_2 = 32.4$, $p < .001$. There was also a weak interaction between recency and the presence/absence of the connective: $F_1 = 5.9$, $p = .022$, $F_2 = 1.7$, $p = .20$, indicating

that the recency effect was slightly larger when the two names were in different sentences than in two coordinated clauses. Importantly, there was no main effect of clause linkage ($F_1 < 1$; $F_2 = 1.5$, $p = .22$): Overall, there were similar proportions of pronouns in two coordinated clauses and two sentences conditions. Thus, there was no evidence that interference is stronger when the clauses are more tightly linked.

Conclusions

Together, the experiments show that clausal structure affects the choice of anaphoric expression in several ways. First, when clause status is controlled (main-main conditions) or counterbalanced (mean of main-sub and sub-main conditions), language producers choose more pronouns and fewer names for referents in a more recent than a distant clause. This recency effect is slightly weaker when a coordinating connective links the two clauses than when there is no connective. Second, Experiment 3 showed that they also use more pronouns for referents in a main than a subordinate clause. Finally, and probably most interesting, they produce fewer pronouns when the referent is in a clause that is of the same type as another clause in the preceding sentence (both main or both subordinate) than when it is in a clause of a different type. This suggests that two clauses of the same type interfere and thereby reduce the accessibility of the entities that are mentioned in them. As a result, language producers use more explicit expressions. One possibility is that clausal interference occurs if two clauses are syntactically similar. Alternatively, the interference effect may occur because of similarity in information structure: Main clauses contain foregrounded information, whereas subordinated clauses contain backgrounded information (Levinson, 1983; Matthiessen & Thompson, 1988; Miltsakaki, 2002). In either case, our experiments provide evidence that interference during reference production not only occurs due to similarity of semantic properties between the referent and competitors, but also due to non-semantic properties.

References

- Ariel, M. (1990). *Accessing noun-phrase antecedents*. London: Routledge.
- Arnold, J. E. (2008). Reference production: Production-internal and addressee-oriented processes. *Language and Cognitive Processes*, 23, 495–527.
- Arnold, J. E., & Griffin, Z. M. (2007). The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language*, 56, 521–536.
- Clancy, P. M. (1980). Referential choice in English and Japanese narrative discourse. In W. L. Chafe (Ed.), *The pear stories: Cognitive, cultural and linguistic aspects of narrative production* (pp. 127–202). Norwood, NJ: Ablex.
- Cooreman, A. & Sanford, A. J. (1996). *Focus and syntactic coordination in discourse*. Edinburgh: HCRC publications.

- Fletcher, C. R. (1984). Markedness and topic continuity in discourse processing. *Journal of Verbal Learning and Verbal Behavior*, 23, 487–493.
- Fukumura, K., Hyönä, J., & Scholfield, M. (in press). Gender affects semantic competition: the effect of gender in a non-gender marking language. *Journal of Experimental Psychology: Learning, Memory and Cognition*.
- Fukumura, K., & Van Gompel, R. P. G. (2010). Choosing anaphoric expressions: Do people take into account likelihood of reference? *Journal of Memory and Language*, 62, 52–66.
- Fukumura, K. & Van Gompel, R. P. G. (2011). The effect of animacy on the choice of referring expression. *Language and Cognitive Processes*, 26, 1472–1504.
- Fukumura, K., Van Gompel, R. P. G., Harley, T., & Pickering, M.J. (2011). How does similarity-based interference affect the choice of referring expression? *Journal of Memory and Language*, 65, 331–344.
- Fukumura, K., Van Gompel, R. P. G., & Pickering, M. J. (2010). The use of visual context during the production of referring expressions. *Quarterly Journal of Experimental Psychology*, 63, 1700–1715.
- Givón, T. (1983). Topic continuity in discourse: An introduction. In T. Givón (Ed.), *Topic continuity in discourse: A quantitative cross-language study* (pp. 1–41). Amsterdam: John Benjamins.
- Grosz, B. J., Joshi, A. K., & Weinstein, S. (1995). Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics*, 21, 203–225.
- Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of anaphoric expressions in discourse. *Language*, 69, 274–307.
- Harris, C. L. & Bates, E. A. (2002). Clausal backgrounding and prominental reference: A functionalist approach to c-command. *Language and Cognitive Processes*, 17, 237–269.
- Levinson, Stephen C. (1983). *Pragmatics*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.
- Matthiessen, C. & Thompson, S. A. (1988). The structure of discourse and 'subordination'. In J. Haiman & S. A. Thompson (Eds.), *Clause combining in grammar and discourse language* (pp. 275–329). Amsterdam: John Benjamins.
- Miltsakaki (2002). Toward an aposynthesis of topic continuity and intrasentential anaphora. *Computational Linguistics*, 28, 319–355.
- Morrow, D. G. (1985). Prominent characters and events organize narrative understanding. *Journal of Memory and Language*, 24, 304–319.
- Price, J. M. (2008) *The use of focus cues in healthy ageing*. PhD thesis, University of Glasgow.
- Stevenson, R. J., Crawley, R. A., & Kleinman, D. (1994). Thematic roles, focus and the representation of events. *Language and Cognitive Processes*, 9, 519–548.