Journal of Memory and Language 66 (2012) 479–508

Contents lists available at SciVerse ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Discourse expectations and relative clause processing

Douglas Roland [a,*], Gail Mauner [b], Carolyn O'Meara [a,1], Hongoak Yun [b]

[a] Department of Linguistics, University at Buffalo, The State University of New York, Buffalo, NY 14260, United States
[b] Department of Psychology, University at Buffalo, The State University of New York, Buffalo, NY 14260, United States

ARTICLE INFO

Article history:
Received 27 February 2011
revision received 4 December 2011
Available online 21 January 2012

Keywords:
Sentence processing
Relative clause
Discourse
Constrain-based approach
Corpus analysis

ABSTRACT

We investigated the role of discourse context in relative clause processing. We first replicated Reali and Christiansen's (2007a) finding that pronominal object relative clauses are easier to process than analogous subject relative clauses (an effect which stands in contrast to previous research on pronominal relative clauses). We then analyzed corpus data to demonstrate that the embedded noun phrase in object relative clauses nearly always has a discourse-old referent, even if the noun phrase is not pronominal, while the referent of the embedded noun phrase in subject relative clauses is typically discourse-new. We then extended the work of Mak et al. (2008), by demonstrating in region-by-region reading that full noun phrase object relative clauses are not more difficult to process than subject relatives when they are preceded by appropriate discourse contexts. Finally, we reanalyzed data from the Dundee Eye-tracking corpus to show that contra Demberg and Keller (2007), naturally occurring object relatives are no harder to process than subject relatives. We conclude that the processing difficulties associated with object as compared to subject relative clauses arises because object relative clauses violate more discourse expectations in typical experimental contexts.

Introduction

The experimental investigation of language comprehension is typically carried out by examining the processing of isolated sentences. This is a near necessity due to need to exert experimental control over extraneous variables. However, the use of isolated contexts can produce misleading results. If the comprehender has expectations associated with a structure based on the contexts in which the structure normally appears or the discourse functions that the structure normally serves, these expectations may be violated when the structure appears in an isolated context. Furthermore, if two structures have differing degrees of unnaturalness in isolation, and they are compared in isolated contexts, the results of the comparison may be misleading, as the differing effects of the isolated contexts will be confounded with the experimental variables being manipulated.

We argue that the well-known differences between the degree of difficulty associated with the processing of subject and object relative clauses are at least in part due to the differing degrees to which subject and object relative clauses violate expectations when encountered in isolated contexts. Using a combination of corpus data, reading time experiments, and statistical modeling of data from an eye-tracking corpus, we explore the expectations associated with subject and object relative clauses and investigate the extent to which these expectations influence processing. This allows us to re-evaluate the extent to which previous results support various models of processing.

In general, object relative clauses such as (1) have been found to be more difficult to process than analogous subject relative clauses, such as (2). However, a recent series of findings demonstrate that object relative clauses are not always more difficult to process than subject relative

* Corresponding author.
E-mail address: droland@buffalo.edu (D. Roland).
[1] Present address: Seminario de Lenguas Indígenas, Instituto de Investigaciones Filológicas, Universidad Nacional Autónoma de México, Mexico.

0749-596X/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2011.12.004

clauses. Rather, in some situations, they are easier to process than analogous subject relative clauses. The difficulty in processing object relative clauses is either reduced or eliminated in examples like (3), where the modified noun phrase is inanimate (Gennari & MacDonald, 2008; Traxler, Morris, & Seely, 2002, inter alia). Additionally, pronominal object relative clauses like (4) are read more quickly than analogous subject relative clauses like (5) (Reali & Christiansen, 2007a). Finally, Mak, Vonk, and Schriefers (2008) demonstrated that the difficulty in processing Dutch object relative clauses can be eliminated by changing the ongoing topic of the context which precedes the relative clause (see discussion below for examples).

(1) The lady that the banker visited enjoyed the dinner very much.
(2) The lady that visited the banker enjoyed the dinner very much.
(3) The movie that the director watched received a prize.
(4) The lady that you visited enjoyed the dinner very much.
(5) The lady that visited you enjoyed the dinner very much.

The discourse functions of object relative clauses described by Fox and Thompson (1990) have been proposed as either a direct or indirect possible cause for the reduction in difficulty in all three cases where object relative clause difficulty is reduced or eliminated (Gennari & MacDonald, 2009; Mak et al., 2008; Reali & Christiansen, 2007a). Fox and Thompson studied the use of relative clauses in a corpus of spoken conversational English. They state that a referent must be relevant for a listener at the point at which it is introduced. The main way in which a referent is made relevant is by relating it to another referent whose relevance is clear – a process known as *grounding*. One way in which grounding is accomplished is by modifying the ungrounded referent with a relative clause, where the embedded noun phrase refers back to a relevant referent in the ongoing discourse (e.g., if the ongoing topic is the speaker's pet dog, and a squirrel is suddenly mentioned, the referent of the squirrel can be grounded by saying *The squirrel that he chased*...). Because the embedded noun phrase refers back to a referent in the ongoing discourse, it tends to be a pronoun – resulting in a pronominal relative clause. For reasons discussed below, this process also tends to result in object relative clauses more often than subject relative clauses. This leads to a distributional difference in the occurrence of pronouns and full noun phrases in subject and object relative clauses, with subject relative clauses tending to have full noun phrases, and object relative clauses tending to have pronominal noun phrases.

Reali and Christiansen attribute their finding that pronominal object relatives are read more quickly than pronominal subject relatives to the relative frequencies of *that + pronoun + verb* vs. *that + verb + pronoun* and *that + full noun phrase + verb* vs. *that + verb + full noun phrase*. They suggest that word chunk frequencies and the discourse factors outlined by Fox and Thompson combine with other factors such as memory limitations to influence relative clause processing.

Similarly, the processing differences between object relative clauses modifying animate nouns and those modifying inanimate nouns may also be due to frequency differences attributable to the discourse functions described by Fox and Thompson. Fox and Thompson point out that because inanimate referents are more likely to be ungrounded than animate referents, they are more likely to be modified by relative clauses. Moreover, when the modified noun phrase is non-human/inanimate, the relative clause usually takes the form of an object relative clause. This is because the ongoing topics of conversation that the embedded noun phrases in relative clauses refer to tend to be participants in the conversation (or at least other human/animate entities). Given a relationship between an inanimate object (the referent being modified), and a human referent, it is more likely for the human to be doing something to the inanimate referent (resulting in an object relative clause such as *the clay that the potter threw*) than for the inanimate referent to be doing something to the human (resulting in a subject relative clause such as *the clay that annoyed the potter*). This results in a distributional difference between object and subject relative clauses depending on whether they modify animate or inanimate noun phrases, with inanimate noun phrases being more likely to be modified by object relative clauses, and animate noun phrases being more likely to be modified by subject relative clauses. This distributional difference is reflected in sentence completions (Gennari & MacDonald, 2008), where participants were more likely to complete a prompt with an animate subject (e.g., *the musician that*...) with a subject relative clause, and to complete a prompt with an inanimate subject (e.g., *the movie that*...) with an object relative clause. This distributional difference has been suggested as a cause for the processing differences between relative clauses modifying animate noun phrases and those modifying inanimate noun phrases (e.g., Gennari & MacDonald, 2008, 2009; Mak, Vonk, & Schriefers, 2006).

The discourse function of object relative clauses described by Fox and Thompson also appears to underlie recent results in Dutch relative clause processing (Mak et al., 2008). Dutch relative clauses are different from English relative clauses in that Dutch relative clauses are structurally ambiguous. In both subject (6) and object (7) relative clauses, the modified noun phrase is followed by the relativizer, and then the embedded noun phrase. Unless the modified noun phrase is a case-marked pronoun, it is unclear whether the noun phrase is an object or a subject, and thus, whether the relative clause is a subject relative clause or an object relative clause. In a sentence starting with *De politie heeft de bewoners, die de inbreker* ..., the phrase *die de inbreker* could either be the start of a subject relative clause, as in (6), where *de inbreker* is ultimately the object of *neergeslagen*, or an object relative clause, as in (7), where *de inbreker* is ultimately the subject of *neergeslagen*. The initial interpretation depends on the relative likelihood of *de bewoners* (*the occupants*) and *de*