

## Investigating effects of structural and information-structural factors on pronoun resolution

ELSI KAISER

### 14.1 Introduction

The ability to interpret informationally impoverished referential forms (e.g. *s/he, it*) is crucial for successful language comprehension. However, this task appears rather daunting, given that a form such as *she* can, at least in principle, refer to any female human referent—in other words, any member of a rather sizeable set. The “search space” is commonly assumed to be constrained by the nature of the mapping between referential forms and referent salience. Specifically, it is widely agreed that there exists a connection between the degree of reduction of referential forms and the level of salience of their antecedents, such that the most reduced (and semantically least informative) forms refer to the most salient referents—in other words, to the most prominent, most accessible referents. A range of factors has been argued to influence salience and to make referents good antecedents for subsequent pronouns, but questions remain regarding the nature of the interaction between the different factors.

This chapter explores two questions related to reference resolution: (i) How do information-structural factors and syntactic factors (in particular focusing, pronominalization, and subjecthood) interact to guide the process of

I would like to thank Rebekka Puderbaugh, Sasha Eloi, and Joyce McDonough for assistance with the experiment described here. I would also like to thank Christine Gumlogson, Jeffrey Runner, and Michael Tanenhaus for useful feedback and comments on earlier versions of some aspects of this chapter. Thanks also to audiences at *Sinn und Bedeutung 10* (Berlin 2005) and the 81st Annual Meeting of the LSA (Anaheim 2007) where some parts of this research were presented. Many thanks are also due to Caroline Féry and Thomas Weskott for their insightful comments. Needless to say, all mistakes and misunderstandings are mine.

reference resolution? (ii) Is the interpretation of a sentence-initial anaphoric form connected to the presence or absence of other referential forms later in the same sentence? The first question looks backwards from an anaphoric form with the goal of clarifying what aspects of earlier discourse influence pronoun interpretation. In contrast, the second question looks forward to investigate whether upcoming discourse has an impact on reference resolution.

The results of the sentence-completion experiment discussed in this chapter are compatible with the view that multiple, differently weighted factors influence reference resolution (see Kaiser and Trueswell 2008 and others). The results indicate that a subject-position pronoun is most likely to be interpreted as referring to the agentive subject of the preceding sentence, regardless of whether the subject is discourse-old and pronominalized (e.g. *He criticized the businessman. He . . .*) or discourse-new and contrastively focused (e.g. *It was the secretary who scolded her. She . . .*). Assuming that pronouns refer to salient antecedents, this indicates that (agentive) entities realized in the grammatical role of subject can be highly salient even when they are not discourse-old or pronominalized (what some might call “topical”). The results also provide some indication that although clefting a focused object makes it more likely to be the antecedent of a subsequent pronoun, focusing may contribute less to a referent’s salience than subjecthood and pronominalization.

As for the effects of subsequent discourse, the results support the view that a purely backward-looking approach to reference resolution is insufficient (see also Centering Theory, Grosz, Joshi, and Weinstein 1995). More specifically, the data show that whether an ambiguous pronoun in subject position refers to the preceding subject or the preceding object is connected to the referential properties of the remainder of the clause—specifically, whether or not the other argument from the preceding sentence is mentioned in the pronoun-containing clause. A pronoun that refers to the *preceding subject* is more likely to be followed by a subsequent mention of the preceding object than a pronoun that refers to the *preceding object* is to be followed by a mention of the preceding subject. This pattern suggests that if a less salient referent (object) is “upgraded” by being interpreted as the antecedent of a subject-position pronoun, then subsequent mention of the higher-salience referent (subject) in the same clause is avoided.

As a whole, the analyses presented in this chapter show that subjecthood plays an important role both when looking back at preceding discourse and when looking forward to subsequent discourse. The structure of the chapter is as follows. Section 14.2 provides an overview of the notion of salience and discusses some factors argued to influence referent salience. In Section 14.3, I explore the idea—discussed in Centering Theory (e.g. Grosz, Joshi, and

Weinstein 1995)—that the interpretation of a pronoun in subject position depends on or has consequences for what comes later in the same sentence (see also Winograd 1972, Hobbs 1979, Kehler 2002 for relevant discussion). Section 14.4 presents the results of the sentence completion experiment, and conclusions and directions for future work are discussed in Section 14.5.

## 14.2 Saliency

It is commonly assumed in the tradition of attention-based approaches to reference resolution that use of different kinds of referential expressions (pronouns, demonstratives, full NPs, etc.) is correlated with the level of saliency/accessibility of their antecedents (e.g. Givón 1983, Ariel 1990, Gundel, Hedberg, and Zacharski 1993). Generally, researchers assume that the most reduced anaphoric expressions refer to the most salient antecedents, where “salient” is used to mean entities that are currently most prominent, at the center of attention. If we accept this view, we can use pronouns as a tool to investigate saliency: we can probe what factors make an entity likely to be the antecedent of a pronoun in subsequent discourse, and thus learn more about what factors influence saliency.<sup>1</sup> Previous research has argued that a range of factors increase the likelihood of subsequent pronominalization (see Arnold 1998 for a summary), including (i) occupying the grammatical position of subject, (ii) being given/old information, and (iii) being realized as a pronoun.

Many of the factors claimed to affect subsequent pronominalization seem to be linked with the notion of topicality. A precise definition of topicality, however, has proved hard to pin down, and the term is used differently by different researchers. According to Reinhart’s aboutness-based definition, the topic is “the expression whose referent the sentence is about” (Reinhart 1982: 5). Gundel (1985) characterizes topics in terms of shared knowledge: “the topic of a speech act will normally be some entity that is already familiar to both speaker and addressee” (Gundel 1985: 92). Strawson defines the topic of an utterance as “what is of current interest or concern” (Strawson 1964: 104).

<sup>1</sup> Assuming that referential forms can be mapped directly onto a saliency scale is not entirely unproblematic (see Kaiser 2003, Kaiser and Trueswell 2008). Our research suggests that not all referential forms are sensitive to the same factors to the same degree, and thus the mapping between referential forms and degree of saliency of the antecedent may well be more complex than is normally assumed. However, this chapter focuses only on one referential form, namely, personal pronouns in English, and our conclusions should be interpreted as restricted to this form. For present purposes, we assume that English personal pronouns can be used as a tool to probe referent saliency. However, even if one regards this saliency-related assumption as problematic, the results are nevertheless informative, as they shed light on the factors influencing pronoun interpretation.

In recent Centering-theoretic work, E. F. Prince (2003) and Beaver (2004) use the term “topic” for the backward-looking center in Centering Theory (Grosz *et al.* 1995)—a use which links topicality with pronominalization and givenness. Although it is not the aim of this chapter to provide a precise definition of topicality, it is worth noting that many of the factors that have been claimed to influence saliency have also been linked to the notion of topicality.

However, not all researchers agree that topicality-related factors are the primary determinants of referent saliency, and some have suggested that focusing is what makes referents cognitively prominent. In the remainder of this section, we examine the claims regarding subjecthood, givenness, and pronominalization, which could be regarded as being related to topicality, and then turn to research on focusing.

### 14.2.1 Effects of subjecthood, givenness, and pronominalization

One of the most widely discussed saliency-influencing factors is grammatical role. A sizeable body of research suggests that entities realized in subject position are more salient than those in non-subject positions (e.g. Chafe 1976, Brennan, Friedman, and Pollard 1987, Crawley and Stevenson 1990; see also Gordon, Grosz, and Gilliom 1993, *inter alia*), a claim that has received support from corpus work and psycholinguistic research (see also Fiedler *et al.*, in this volume, for typological discussion concerning the special status of subjects). (We focus here on pronouns in subject position. The interpretation of object-position pronouns has also been investigated in previous work, e.g. Smyth 1994, Chambers and Smyth 1998.)

Another factor that has been linked to saliency is the distinction between given and new information. On the basis of a detailed corpus analysis of a range of German texts, Strube and Hahn (1996, see also Strube and Hahn 1999 for further details) claim that old/given referents are more salient than new referents. According to their view, an anaphoric expression that follows a sentence containing a discourse-old referent and a discourse-new referent tends to refer to the discourse-old one (see also Ariel 1990 and Givón 1983 on topicality effects). Further evidence regarding givenness comes from corpus work on the Oceanic language Japanese by Ballantyne (2004).

A further factor that has been argued to influence saliency is the referential form with which an entity is realized. According to Kameyama (1996), the saliency of a referent in non-subject position increases if that referent is pronominalized, to the extent that the referent can “compete” with a non-pronominalized entity in subject position (see also Grosz, Joshi, and Weinstein 1986). In Beaver’s (2004) Optimality-theoretic approach to

anaphora resolution, the constraint SALIENT FORM similarly “implies that being pronominalized makes a referent salient in the future” (Beaver 2004: 31 n. 30; see also Kehler 2002: 169).

Given that factors such as subjecthood, givenness, and pronominalization have been claimed to increase salience, and that these factors are often regarded as related to topicality, one might be tempted to conclude that a salient referent is a topical referent. However, not all researchers agree that salience is connected to topicality, as the next section makes clear.

#### 14.2.2 *Effects of focusing*

This section examines claims that focusing increases referent salience—more specifically, that contrastive focus increases salience. The existing psycholinguistic work on this topic has, for the most part, used structures such as *it*-clefts (*It was Mary who called Lisa*), which express contrastive/identificational focus.<sup>2</sup> In keeping with the existing emphasis on contrastive focus, the experiment described in Section 14.4 uses *it*-clefts and *in-situ* focus constructions<sup>3</sup> in contrastive contexts, and thus the conclusions drawn on the basis of this experiment apply only to contrastive focus.

A large body of cognitive psychology research indicates that focused entities pattern differently from non-focused entities. In an early study, Hornby (1974) presented participants with pictures and sentences and asked them whether the sentence matches the picture. When presented with cleft sentences (e.g. *It is the girl who is riding the bicycle*), participants were better at detecting mismatches when the mismatching information was focused than when it was presupposed (see also Cutler and Fodor 1979). Related work suggests that people attend more to the focused/non-presupposed part of a sentence (Zimmer and Engelkamp 1981, but see Delin 1990), and remember focused information better than non-focused information (e.g. Singer 1976, Birch and Garnsey 1995; see also Birch and Rayner 1997).

Extending the investigation of focusing effects to reference resolution, Almor (1999) showed that definite NPs referring to focused antecedents are read faster than definite NPs whose antecedents are not in focus, regardless of whether the focused entity is the subject or object. His results suggest that reference to focused antecedents is processed more easily than reference

<sup>2</sup> Corpus studies show that regarding all clefts as structures where the clefted constituent is contrastively focused and the rest of the sentence is presupposed is an oversimplification (see e.g. Delin 1990). Nevertheless, the *it*-clefts used in the experiment described here were straightforward in that the focused constituent was contrastively focused new information and the rest of the sentence was old/presupposed information.

<sup>3</sup> Green and Jägger (2003) argue that *in-situ* focus in English can be interpreted contrastively, contrary to É. Kiss (1998).

to non-focused antecedents. Further evidence for effects of focusing comes from Birch, Albrecht and Myers (2000), who argue, based on sentence completion results, that focused concepts are more salient and have a “stronger memory trace” (Birch *et al.* 2000: 302) than non-focused concepts. More recently, Foraker and McElree (2007) claimed that clefting makes referents more distinctive in memory but does not cause them to be actively maintained in the focus of attention. Their results fit with the other studies insofar as they suggest that focused entities are represented differently from non-focused entities.

In sum, the existing literature leaves us at an intriguing crossroads. A number of factors—including subjecthood, givenness, and pronominalization—are claimed to increase the discourse salience of an entity and thus make that entity a good antecedent for a subsequent pronoun, and many of these factors seem to be related to topicality. However, a number of other studies have found that contrastive focus affects the discourse representation of an entity, rendering it more prominent in people’s mental models of the discourse.

#### 14.2.3 *Comparing focus to other factors*

Taking important steps to reconcile the claims discussed in the preceding sections, Arnold (1999), Cowles (2003), and Cowles, Walenski, and Klueder (2007) investigated how focus, subjecthood, and givenness influence salience. Because these experiments provide the starting point for much of the research presented in this chapter, this section summarizes their key findings.

Given that not all researchers agree on how the term “topic” should be used and that this chapter does not aim to provide a definition of topichood, I will adapt a neutral approach and describe the referents in Arnold’s and Cowles’s experiments in terms of (i) their grammatical role, (ii) their referential form (full NP vs. pronoun), and (iii) their discourse status (given/old vs. new). In adopting this approach, I do not intend to suggest that topicality is unimportant. Whether an entity is realized in subject or object position, realized with a pronoun or a full NP, and so on, depends on and has consequences for higher-level issues connected to aboutness and topicality that merit further research (see also Fiedler *et al.*, this volume, for related discussion concerning the subject position and topicality).

In a series of experiments, Arnold (1999, see also Arnold 1998) used pronouns as a tool for probing which referent in the preceding discourse is the most salient, and how salience is influenced by topicality and contrastive focus. On the basis of a rating study comparing focused (clefted) objects and matrix subjects (which she analyzes as sentence topics, following a large

body of research on the connection between subjecthood and topicality) to other referents, Arnold concluded that topics and foci are more salient than other referents. In two further experiments, she pitted pronominalized, discourse-old subjects against focused objects, and discourse-new subjects against focused objects. She concluded that (i) pronominalized, discourse-old subjects—which she views as discourse-topics—are more salient than focused objects (both *in-situ* and clefted), but that (ii) discourse-new subjects are relatively less salient than discourse-old subjects. She suggests that a focused object is better able to compete with a discourse-new full-NP subject than with a discourse-old pronominalized subject, which she regards as an established discourse-topic. She views salience as a competitive phenomenon, “where the representations of different referents in a particular discourse compete for activation” (Arnold 1999: 28).

In related research, Cowles (2003, see also Cowles *et al.* 2007) reached a somewhat different conclusion concerning the effects of topicality and focus. Cowles tested three types of antecedent: (i) focused subjects, (ii) discourse-old full NP subjects, and (iii) discourse-new full NP subjects. On the basis of a cross-modal naming study, she concluded that “[a]ll three information statuses [discourse-topic, sentence-topic and contrastive focus, EK] appear to make their referent more likely to be interpreted as the antecedent of a subsequent pronoun” (Cowles 2003: 93). Thus, in contrast to Arnold, who found that established discourse-topics (pronominalized subjects) are more salient than focused objects, Cowles’s results lead her to conclude that “two information-structure types [topic and focus, EK] that are considered distinct... appear to have the same psychological effect” (Cowles 2003: 94). However, as Cowles *et al.* (2007) note, detailed analysis of the results hints that focusing may not have quite as powerful an effect as subjecthood does, although they note that this is not entirely clear.

The differences in Arnold’s and Cowles’s findings—in particular Arnold’s conclusion that (discourse) topics are more salient than foci and Cowles’s conclusion that topics and foci have comparable effects on salience—may stem from methodological differences. However, the disparity could also stem from differences in grammatical role and pronominalization. Arnold’s finding that discourse-old pronominalized subjects (which she analyzes as discourse-topics) are more prominent than focused objects could be related to their respective grammatical roles (and/or NP form) in addition to their topic/focus status. In Cowles’s experiments, both topics and foci were full names in subject position, which may have meant that (i) Cowles’s foci (in subject position) were more prominent than Arnold’s foci (in object position), due to the difference in grammatical role and/or that (ii) Cowles’s topics (full NPs in

subject position) were less prominent than Arnold’s pronominalized subjects, due to effects of NP form.

Thus, although existing experimental work suggests that focused entities and discourse-old subjects (which some analyze as topics) are more salient than other referents (see also Navarretta 2002), questions remain as to how the different factors involved (focus, subjecthood, pronominalization/givenness) relate to each other. The experiment reported here builds on Arnold’s and Cowles’s work by investigating both focused subjects and objects, as well as pronominalized subjects and objects.

#### 14.3 Connection between pronoun interpretation and mention of other referents

The work on reference resolution discussed in the preceding sections focused primarily on what aspects of the preceding discourse influence the interpretation of pronouns in subject position. In this section I consider the role of subsequent discourse, namely, whether the interpretation of a subject pronoun is connected to the presence of other referential forms later in the same sentence. It is clear that interpretation of a subject-position pronoun can be influenced by the semantics of the verb that follows the pronoun (see e.g. Winograd 1972), but here we focus on the potential effects of another kind of post-pronominal information, namely the referential properties of the rest of the clause.

The idea that the interpretation of the initial (subject) pronoun is connected to the likelihood that another referent will or will not be mentioned is addressed in Centering Theory, an influential computationally based model of the local-level component of attentional state in discourse. In this section, I first review some basic tenets of Centering Theory, and then in Section 14.3.1 I outline how, according to Centering, the presence or absence of other referents in the pronoun-containing clause is connected to the interpretation of the subject-position pronoun.

Centering Theory, formulated on the basis of extensive corpus work, makes predictions regarding how focus of attention, choice of referring expression and local coherence are connected (Grosz, Joshi, and Weinstein 1995). According to Centering, entities mentioned in an utterance (called “centers”) are ranked in terms of how salient (how “central”) they are. It is commonly assumed that in English, this ranking is done by grammatical role, with subjects being more salient than objects. Centering Theory provides an implementable way of analyzing the coherence relations between pairs of utterances—in other words, of assessing how much of a load they impose on the human language processing system. The coherence of the transition

from one utterance to the next utterance depends (i) on whether, (ii) in what position, and (iii) with what form the most central entity from one sentence is mentioned in the next sentence. For example, using a subject pronoun in Utterance 2 to refer to the most central entity mentioned in Utterance 1 results in a more coherent transition than use of an object pronoun in Utterance 2 to refer to the most central entity in Utterance 1. Centering Theory posits four transition types, of differing levels of coherence (i.e. involving different amounts of processing load), summarized in the Appendix.<sup>4</sup>

Although the original Centering-theoretic work by Grosz *et al.* (1995 and earlier versions) is not a pronoun-resolution algorithm, an algorithm for pronoun interpretation based on Centering principles was proposed by Brennan, Friedman, and Pollard (1987, BFP). The BFP algorithm assumes that more coherent transitions are preferred over less coherent transitions (i.e. there is a bias to minimize processing load), and the algorithm assigns pronominal reference accordingly (see also Poesio *et al.* 2004). Thus, pronouns are resolved in such a way that the resulting transition is as coherent as possible.

#### 14.3.1 *Effects of subsequent referents on pronoun interpretation*

To illustrate how, in Centering Theory, the interpretation of the initial pronoun is connected to the likelihood of subsequent mentions of other entities, let us consider (1). I start by considering a situation where the subject pronoun at the start of the third sentence is interpreted as referring to the preceding object:

- (1) a. John was sitting by the window. Peter tickled him<sub>John</sub>. He<sub>John</sub> laughed.  
 b. John was sitting by the window. Peter tickled him<sub>John</sub>. He<sub>John</sub> tickled {him<sub>Peter</sub>/Peter} back.

According to Centering Theory, in a context such as (1), if the subject pronoun *he* in the third sentence is interpreted as referring to the preceding object (*John*), then continuation (1a)—in which the preceding subject Peter is not mentioned—is more coherent than continuation (1b), in which the preceding subject Peter is mentioned in a non-subject position later in the sentence (see Kehler 1997 for discussion of similar examples).

Centering Theory allows for the possibility that not only grammatical role but also different syntactic constructions (e.g. clefts) and semantic/pragmatic

<sup>4</sup> The reader is referred to Grosz *et al.* (1995) and Walker, Joshi, and Prince (1998) and the references cited therein for further details (see also Poesio, Stevenson, di Eugenio, and Hitzeman 2004 for recent and extensive Centering Theoretic analyses of large corpora).

factors influence which entities are most central. However, for the purposes for this chapter, I make the simplifying assumption that the predictions regarding the consequences of the initial pronoun's interpretation are the same for clefts and SVO sentences. Centering Theory also makes claims regarding the referential form used to refer to the preceding subject (Rule 1, see e.g. Walker *et al.* 1998: 4), which I do not discuss here. In the subsequent discussion, pronominal and full NP mentions are grouped together (as in (1b)), but I acknowledge the importance of distinguishing them in future work.

Let us now consider a situation where the subject pronoun *he* in the third sentence refers to the preceding full NP subject (Peter), as in (1c, d). In this situation, according to Centering Theory, a continuation that mentions the preceding object (1d) is as coherent as one which does not (1c). However, neither is as coherent as continuation (1a), in which the subject pronoun refers to the preceding pronominalized object and no other referents are mentioned.

- (1) c. John was sitting by the window. Peter tickled him<sub>John</sub>. He<sub>Peter</sub> laughed.  
 d. John was sitting by the window. Peter tickled him<sub>John</sub>. He<sub>Peter</sub> pushed {him<sub>John</sub>/John} over.

If we follow Brennan *et al.* and assume that maximally coherent transitions are easiest to process and to be preferred, we expect that in sequences like (1), (i) a subject pronoun referring to the preceding (pronominalized) object will probably not be followed by a mention of the preceding subject later in the sentence (i.e. (1a) is more coherent than (1b)), whereas (ii) a subject pronoun that refers to the preceding subject may or may not be followed by a mention of the preceding object ((1c) is as coherent as (1d)).

As Kehler (1997) notes, the idea that the maximally coherent interpretation of "he" depends on post-pronominal information seems to be at odds with the idea that pronouns in subject position are interpreted quickly and incrementally (e.g. Arnold, Eisenband, Brown-Schmidt, and Trueswell 2000; see Kehler 1997). However, a sizeable body of psycholinguistic research suggests that the human language-processing system can consider different structures/interpretations in parallel (e.g. Tanenhaus and Trueswell 1995 for an overview). This raises the possibility that upon encountering a pronoun, the language comprehension process does not pick a single antecedent but keeps track of ranked alternatives. We could have a system that creates an initial ranking of possible referents incrementally, but is capable of changing the ranking of the alternatives later. If this is on the right track, the claim that interpretation of the sentence-initial pronoun is connected to, or has consequences for, what gets mentioned later seems less implausible.

#### 14.4 Experiment

The experiment discussed in this chapter has two main aims: (i) to investigate how information-structural and syntactic factors (focus, pronominalization, and subjecthood) influence subsequent pronoun use, and (ii) to shed light on whether the interpretation of one referential form (the pronoun in subject position) is connected to the likelihood of other referents being mentioned in subsequent parts of the utterance.

To investigate these questions, I used a sentence-completion task based on dialogues between two hypothetical speakers: speaker A and speaker B. On critical trials, speaker B's response contained the critical target sentence, and ended with a prompt pronoun. Participants played the role of speaker B, and their task was to provide a natural-sounding continuation beginning with the prompt pronoun.

In the critical sentence, syntactic form (cleft vs. SVO) and the grammatical role of the focused constituent<sup>5</sup> were manipulated, as illustrated in (2). This resulted in four conditions: [Cleft.Object = focus], [Cleft.Subject = focus], [SVO.Object = focus] and [SVO.Subject = focus]. In the clefted conditions, the preceding context and the syntactic construction mark one of the arguments as being in contrastive focus (see Rooth 1992); the rest of the sentence is presupposed. In the SVO conditions, the focus-presupposition division is not encoded by a particular syntactic construction but follows from the preceding context. The argument in the presupposed part of the sentence is discourse-old and pronominalized (subject in (2a, c), object in (2b, d)).

- (2) a. [Cleft.Object = focus]  
 A: The maid scolded the bride.  
 B: No, that's wrong! It was the secretary<sub>focus</sub> that she scolded. She ...
- b. [Cleft.Subject = focus]  
 A: The maid scolded the bride.  
 B: No, that's wrong! It was the secretary<sub>focus</sub> who scolded her. She ...
- c. [SVO.Object = focus]  
 A: The maid scolded the bride.  
 B: No, that's wrong! She scolded the secretary<sub>focus</sub>. She ...
- d. [SVO.Subject = focus]  
 A: The maid scolded the bride.  
 B: No, that's wrong! The secretary<sub>focus</sub> scolded her. She ...

<sup>5</sup> Not all factors are fully crossed in this experiment, partly due to the nature of the phenomena under investigation. In this design, pronominalization and focus are in complementary distribution: a particular referent is either pronominalized or focused.

If one follows Reinhart (1982), Givón (1983), Arnold (1999), Cowles (2003) and others in connecting subjecthood with topicality, then it seems to follow that subjects in the [Object = focus] conditions are topics, and subjects in the [Subject = focus] conditions are presumably contrastive topics. However, it seems that some other researchers would disagree with this classification: E.F. Prince (2003) and Beaver (2004) would presumably regard the discourse-old pronominalized object as the topic in the [Subject = focus] conditions, not the focused discourse-new subject. Since a particular definition of topicality is not central for my claims, I will not commit to one view over the other; rather, I will discuss the arguments in terms of their grammatical role, NP form, and focus status.

The verbs used in the 16 target items were agent-patient verbs (as defined by Stevenson, Crawley, and Kleinman 1994); and the three human referents in each target had the same gender; (stereotypically) male or female. The experiment also contained 16 fillers. Participants ( $n = 24$ ) were asked to provide a natural-sounding continuation sentence using the pronoun prompt. They were instructed to imagine that someone had just made the claim in part A, and they were now responding to this person's statement by saying part B out loud and providing a continuation. Parts A and B, as well as the prompt pronoun, were presented to the participants in writing. The responses of the participants were recorded and transcribed. The referent of the prompt pronoun was coded by two coders working independently. If it was not clear who the pronoun referred to, the item was coded as "unclear". The continuations were also coded for whether the other argument of the immediately preceding sentence (i.e. whichever argument the pronoun did not refer to) was mentioned. In this chapter I only discuss a simple level of "other argument" coding, one that does not take into account referential form or specific grammatical position of the other argument.

##### 14.4.1 Predictions

14.4.1.1 *Predictions regarding the interpretation of subject pronouns* Researchers differ as to whether they view anaphor resolution as being a process that is determined by a single factor or guided by multiple constraints (e.g. Ariel 1990, Arnold 1998, Strube and Hahn 1999, Kaiser and Trueswell 2008.). If a single factor determines reference resolution, the predictions are straightforward. If subjecthood is what matters, the prediction is that the prompt pronoun will consistently refer to the immediately preceding subject, regardless of NP form or focus status. (Note that the subjects in the experiment are agentive so the design cannot distinguish subjecthood from agentivity.) If pronominalization

TABLE 14.1 Predicted effects of subject-pronoun interpretation on subsequent mentions

Antecedent of pronoun	Is other referent mentioned later in the sentence?
[Subject = focus] Preceding subject	Yes or no
Preceding object	No
[Object = focus] Preceding subject	Yes or no
Preceding object	(Yes or no)

is crucial, pronouns should always refer to the pronominalized argument in the preceding sentence. If focusing is what matters, prompt pronouns are predicted to refer to the focused referent, regardless of grammatical role.

Turning now to the predictions of a multiple-factor view, if we assume that all three factors contribute equally, we predict an equally strong preference for the subject of the immediately preceding clause in all four conditions. Let us first consider the [Subject = focus] conditions. In both the [SVO.Subject = focus] and the [Cleft.Subject = focus] conditions, pronominalization pushes towards the object, but subjecthood and focusing both point to the subject. Thus, if the three factors contribute equally, the subject is predicted to be the preferred antecedent. Turning now to the [Object = focus] conditions, in both the [SVO.Object = focus] and the [Cleft.Object = focus] conditions, subjecthood and pronominalization point towards the subject, whereas focusing points to the object. Again, if all factors matter equally, the subject is predicted to be the preferred antecedent.

However, it may well turn out to be the case that the three factors are not all weighted equally. In fact, constraint-based models of language processing (e.g. MacDonald, Pearlmuter, and Seidenberg 1994, Trueswell, Tanenhaus, and Garnsey 1994, Tanenhaus and Trueswell 1995; see also Bates and MacWhinney 1989) claim that language processing is guided by constraints of differing weights (see also Arnold 1998). I return to this issue in the results section.

**14.4.1.2 Predictions regarding initial pronoun interpretation and subsequent mentions** The Centering-Theory-derived predictions for how the interpretation of the initial pronoun is connected to the likelihood of the other referent being mentioned in the remainder of the sentence are shown in Table 14.1. These predictions are largely derived from the BFP algorithm, which rests on the assumptions articulated in Centering Theory such as the ranking of the different types of transitions that hold between utterances (see Section 14.3). As mentioned above, for ease of exposition I assume that the predictions are

the same for clefted and non-clefted sentences. In addition, I only discuss a simple level of "other referent" coding that does not take into account the other referent's referential form or grammatical position. In this regard, the data analysis presented in this chapter is best regarded as a simplified adaptation of Centering Theory.

In conditions where the subject is focused and the object is pronominalized (e.g. *The secretary<sub>focus</sub> scolded her. She...*), the BFP algorithm predicts a strong asymmetry between subject interpretations and object interpretations of the sentence-initial pronoun prompt. In subject-focus conditions, if the prompt pronoun refers to the preceding object, the prediction is that the preceding object may or may not be mentioned later in the prompt-pronoun sentence ("yes or no" in Table 14.1). This is because the resulting transitions<sup>6</sup> are equally coherent in Centering terms (both are "smooth-shift" transitions). However, if the prompt pronoun refers to the preceding object, the prediction is that the preceding subject is not mentioned. The transition resulting from not mentioning the subject ("continue", a very coherent transition) and the transition resulting from mentioning the subject ("rough shift", a very incoherent transition) are two extremes of the Centering transition hierarchy, and thus if we follow BFP, we predict a preference for the most coherent transition over the least coherent transition and thus a preference for the subject to not be mentioned.

In conditions where the object is focused and the subject is pronominalized (e.g. *She scolded the secretary<sub>focus</sub>. She...*), we do not expect such a clear asymmetry between subject interpretations and object interpretations. If the prompt pronoun is used to refer to the preceding subject (e.g. *She scolded the secretary. She was really angry.*), the prediction is that the preceding object (*secretary*) may or may not be mentioned. This is because both situations lead to equally coherent transitions. (Both involve a continue transition.) However, if in these conditions the pronoun refers to the preceding (focused) object (e.g. *She (maid) scolded the secretary. She really didn't like being scolded by the maid.*), then according to the traditional view of transition rankings and the BFP algorithm, the expectation is that there will be a preference for the preceding subject to be mentioned in the continuation sentence (which results in a retain transition), as compared to being left unsaid (which results in a less coherent smooth-shift transition). However, a number of Centering researchers have expressed doubts regarding the traditional view that smooth-shift transitions really are less coherent than retain transitions (see e.g. Poesio

<sup>6</sup> It is commonly assumed that four transition types exist, listed here ranked from most coherent to least coherent: Continue > Retain > Smooth-shift > Rough-shift. An overview of some central ideas in Centering Theory is provided in the Appendix, and further details are available in Grosz *et al.* (1995) and Walker *et al.* (1998).

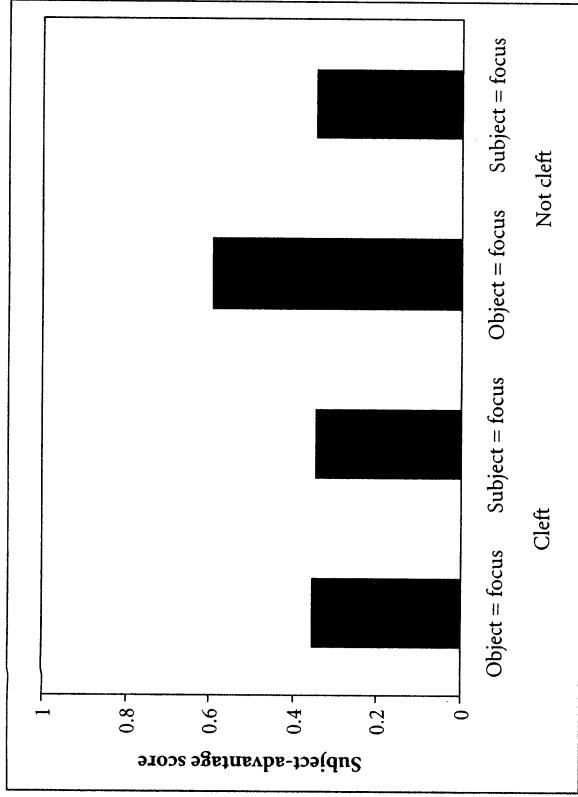


FIGURE 14.1 Subject-advantage scores (proportion of subject continuations minus proportion of object continuations)

*et al.* 2004), and thus a more reasonable prediction might be that if the initial pronoun refers to the preceding object, the subject may or may not be mentioned.

#### 14.4.2 Results and discussion

**14.4.2.1 Interpretation of subject pronouns** The continuations of the participants show a strong overall bias to interpret a prompt pronoun as referring to the agentive subject of the immediately preceding clause, as illustrated in Figure 14.1. The Figure shows the average subject advantage scores obtained by subtracting the proportion of object continuations from the proportion of subject continuations. A positive number means that there were more subject continuations than object continuations, a negative number, that there were more object continuations than subject continuations. There were significantly more subject interpretations than object interpretations in all four conditions.

The overall subject preference shows that (i) grammatical role makes an important contribution to the salience of focused and pronominalized entities and that (ii) neither focus nor pronominalization outweighs subjecthood as there is no overarching preference for pronouns to refer to either focused or

pronominalized antecedents. Rather, what we see is that pronouns are most likely to be interpreted as referring to (agentive) subjects, regardless of whether the subjects are pronominalized or focused (either in a cleft or *in situ*). This finding highlights the importance of taking grammatical role into account when investigating effects of topicality and focusing, and could be taken as an indication that subjects do not need to be discourse-old or pronominalized in order to be salient.

Although these results attest to the importance of grammatical role, should we conclude that subjecthood is the lone factor that determines pronoun interpretation? A number of reasons argue against such a conclusion. First, a large body of existing work has found reliable effects of other factors in addition to subjecthood (see Section 14.3). Second, as discussed in the pre-conditions, given the set-up of this particular experiment, a subject bias could arise even if multiple factors are underlyingly contributing to referent salience and pronoun interpretation. Thus, an overall subject bias does not provide evidence that only subjecthood matters.

Thirdly, as can be seen in Figure 14.1, the subject advantage does not seem to be equally strong in all four conditions; [SVO.Object = focus] shows a slightly stronger subject advantage than the other three conditions. However, analyses of variance by subjects and by items indicate that the differences between [SVO.Object = focus] and the other conditions are not overwhelming, in some cases reaching only marginal significance or less. This means that we should be careful not to attribute too much importance to the asymmetry visible in Figure 14.1 without further empirical evidence. Nevertheless, the results provide preliminary support for the view that multiple factors guide pronoun interpretation.

In fact, the results for most of the conditions seem to fit well with the view that multiple factors matter (i.e. pronominalization, focusing, and subjecthood all play a role) but also provide some hints that focusing may matter less than subjecthood and pronominalization. In such a situation, the expectation is that pronominalized objects should compete with subjects more than focused objects. In other words, conditions where subjecthood and pronominalization coincide (i.e. the [Object = focus] conditions, [SVO.Object = focus] and [Cleft.Object = focus]) should exhibit a stronger subject preference than conditions where subjecthood and focus coincide (i.e. the [Subject = focus] conditions, [SVO.Subject = focus] and [Cleft.Subject = focus]). This prediction fits with the finding that [Subject = focus] conditions exhibit a numerically weaker subject preference than the [SVO.Object = focus] condition, but raises the question of why the [Cleft.Object = focus] condition does not exhibit as strong a subject



preference as [SVO.Object = focus]. A possible reason for the different behavior of the two [Object = focus] conditions has to do with the fact that the [Cleft.Object = focus] condition is the only condition in which the object linearly precedes the subject (and the subject is in a non-prominent syntactic position). In light of existing claims regarding effects of linear order on salience (e.g. Gernsbacher and Hargreaves 1988, 1992), one might expect the object to receive a boost in this condition and thus to weaken the subject preference (see Arnold 1998, 1999 on salience as a competitive phenomenon). If this interpretation is on the right track, then—although there is no clear main effect of cleft vs. SVO—it suggests that clefting a focused object increases its chances of being referred to by a subsequent pronoun.<sup>7</sup>

In sum, if we attribute a greater influence to pronominalization and grammatical role than to focusing and also take into account the potential object-boosting effects of clefting and/or linear order in the [Cleft.Object = focus] condition, we correctly predict the observed data. However, further research should be conducted to confirm whether focusing is indeed weighted less heavily. Crucially, even on the basis of the current results, it is clear that grammatical role needs to be taken into account when investigating effects of notions such as “topic” and “focus” on salience and pronoun interpretation.

**14.4.2.2. Initial pronoun interpretation and subsequent mentions** Now we turn to the question of whether the interpretation of subject pronouns is influenced by (or influences) what is mentioned later in the same clause. Figure 14.2 shows what percentage of subject and object continuations (i.e. continuations in which the prompt pronoun was interpreted as referring to the preceding subject vs. object) mention the other referent from the immediately preceding clause (collapsing across clefts and SVO sentences).

As can be seen in the figure, the interpretation of the initial pronoun is indeed connected to the likelihood of the other referent being mentioned later in the sentence. Overall, there were fewer mentions of the other referent in those continuations in which the sentence-initial prompt pronoun was used to refer to the preceding object than in those continuations in which the prompt pronoun was used to refer to the subject. Furthermore, more detailed analyses

<sup>7</sup> Parallelism may be an additional factor that one might expect to play a role. According to Smyth (1994), pronouns prefer antecedents that are in the same syntactic position as the pronoun. However, according to Smyth, this parallel-position bias holds only when both sentences have the same global structure and matching thematic roles (e.g. *Peter pushed John. Alex pinched him.*) However, an examination of participants' continuations suggests that the required degree of matching across sentences does not appear to be consistently present, which casts doubt on the idea that parallelism is at work here.

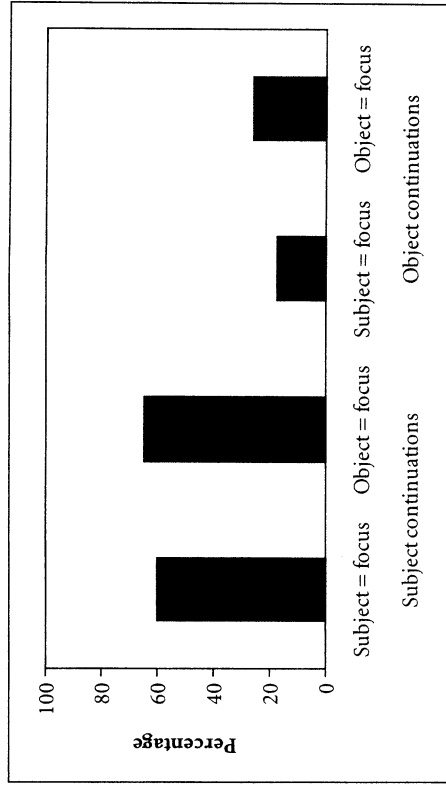


FIGURE 14.2 Percentage of subject continuations and object continuations that mention the other argument from the preceding sentence.

suggest that there do not appear to be overarching effects of topicality, focusing or syntactic form on the likelihood of the other referent being mentioned.

Let us now take a closer look at the [Subject = focus] and [Object = focus] conditions, to see how the data fit with the predictions sketched out in Section 14.4.1.2. In the [Subject = focus] conditions, there is an asymmetry between subject continuations and object continuations, as subject continuations mention the other referent more frequently (> 60%) than object continuations do (< 20%). This fits with the Centering-based predictions in Table 14.1.

In the [Object = focus] conditions, there is an asymmetry in the same direction as with the [Subject = focus] conditions—the other referent seems more likely to be mentioned in subject continuations than in object continuations (approx. 65% vs. < 30%). These data do not fit with the predictions outlined in Section 14.4.1.2, that subject and object continuations would show similar rates of mention, nor do they support the more specific prediction that object continuations would be more likely to involve mention of the other argument than subject continuations.

In sum, the data support the predictions for the [Subject = focus] conditions, but not the predictions for the [Object = focus] conditions. However, although those specific [Object = focus] predictions are not supported, the striking asymmetry between subject and object continuations nevertheless shows that the interpretation of subject-position pronouns is indeed

connected to the likelihood of other referents being mentioned: (i) if the initial pronoun refers to the preceding subject, the continuation often mentions the preceding object as well, but (ii) if the initial pronoun refers to the preceding object, the preceding subject is less likely to be mentioned. Thus, it seems that grammatical role is what matters; there are no clear across-the-board effects of pronominalization, focusing, or syntactic packaging on the likelihood of the other referent being mentioned.

If we assume that subjecthood plays a central role in determining salience, one way of describing the results is as follows: if the sentence-initial pronoun is interpreted as referring to a highly salient referent (the preceding subject), the sentence may also include mention of a lower-salience referent later on. However, if the sentence-initial pronoun is interpreted as referring to a previously lower-salience referent (preceding object), the remainder of the sentence is less likely to make reference to the higher-salience referent of the preceding sentence. In other words, if a less-salient referent is promoted (becomes more salient) by being interpreted as the antecedent for a subject-position pronoun, language users avoid subsequent mention of the preceding higher-salience referent (preceding subject) in a lower-salience position in the same clause. Furthermore, these patterns could also be related to the nature of the semantic relation between the pronoun-initial clause and the preceding clause (e.g. a cause-effect relation or a resemblance relation, see Kehler 2002, Kehler, Kertz, Rohde, and Elman 2008), a question worthy of future investigation.

These findings have implications for our view of reference resolution because they provide further support for the view that pronoun resolution depends on both preceding (pre-pronominal) and subsequent (post-pronominal) information. The data are compatible with an approach which assumes that even though a comprehender can process a sentence incrementally and constructs a weighted set of possible antecedents for a pronoun upon first encountering it, information later on in the sentence can have an impact on the weighting of these alternatives (see also Winograd 1972 on the effects of semantic information).

#### 14.5 Conclusions

This chapter investigated two questions related to reference resolution: (i) How do information-structural factors and syntactic factors (in particular focusing, pronominalization, and subjecthood) interact to guide the process of reference resolution? (ii) Is the interpretation of a sentence-initial anaphoric form connected to the presence or absence of other referential

forms later in the same sentence? The first question looks backwards from an anaphoric form, towards preceding discourse, whereas the second question looks forward, towards subsequent discourse.

When it comes to the first question, the results of the experiment discussed here are compatible with the idea that reference resolution is guided by multiple interacting factors (see also Kaiser and Trueswell 2008 and others). The results indicate that pronouns are most likely to be interpreted as referring to (agentive) subjects regardless of whether they are pronominalized or focused. Participants' continuations indicate that—at least in the contexts investigated here—a contrastively focused, discourse-new, subject is preferred over a discourse-old object. If pronouns are used to refer to highly salient antecedents, our results show that subjects can be highly salient without being discourse-old or pronominalized. In addition, the results provide some hints of (i) object-delts boosting the salience of the focused object (and/or resulting in a loss of salience of the subject) and (ii) a strength asymmetry between pronominalization, subjecthood and focusing, as it seems that pronominalization and subjecthood may contribute more than focusing. These findings, however, are only tentative and need to be investigated more specifically in further work.

The second question sheds light on the question of whether a purely “backward-looking” approach is sufficient for pronoun interpretation. The results support the idea that the interpretation of a pronoun is influenced not only by properties of preceding discourse, but also by what else is mentioned later in the clause containing the pronoun (see Grosz *et al.* 1995, see also Winograd 1972 on effects of semantics). The results show that the interpretation of the subject pronoun (whether it refers to the preceding subject or object) is connected to the likelihood of the other argument from the preceding sentence being mentioned in the same clause: subject continuations were more likely to include a mention of the other argument than object continuations. This finding could be interpreted as follows: If a less salient referent (object) is upgraded by being interpreted the antecedent for a subject-pronoun, subsequent mention of the preceding high-salience referent (preceding subject) in a lower-salience position in the same clause is dispreferred.

Of course, many questions remain open, and further research should be conducted to investigate the validity of the claims discussed here. For example, given data suggesting that different factors may be weighed differently, we would like to know more about the extent of cross-linguistic variation. In addition, the relation between agentivity and subjecthood merits further

study. The experiment discussed here investigated agentive subjects, and thus agentivity and subjecthood are fully correlated. To find out whether the structural notion of subjecthood or the semantics of agentivity is behind the strong effects of subjecthood on subsequent pronominalization, agentive and non-agentive subjects (e.g. experiencers) need to be compared. To learn more about the expectations that participants build up over the course of a sentence after having encountered an anaphoric expression, future work should also test directly the on-line prediction that comprehenders may need to revise their initial preferred interpretation of a pronoun based on subsequent information (see also Kehler 1997). Furthermore, better to understand the nature of post-pronominal effects, it would be useful to investigate the semantic relations between sentences, to see whether different completion types correlate with different inter-sentential relations (e.g. cause-effect vs. similarity, see Kehler 2002, Kehler *et al.* 2008 for details).

In conclusion, the findings presented here fit with the idea that reference resolution is influenced both by preceding discourse and subsequent discourse. Taken as a whole, the results of the two analyses indicate that subjecthood, in particular, plays an important role in guiding both the backward- and forward-looking aspects of pronoun resolution.

### Appendix: Centering Theory

Some of the core ideas, constraints, and rules underlying Centering Theory (from Walker *et al.* 1998: 3–6) are provided here.

- “The set of forward-looking centers,  $Cf(U_i, D)$ , represents discourse entities evoked by an utterance  $U_i$  in a discourse segment  $D$ ” (Walker *et al.* 1998: 3).
- “The highest-ranked member of the set of forward-looking centers is referred to as the preferred center,  $Cp$ ” (Walker *et al.* 1998: 3). The ranking of the  $Cf$  list can depend on various factors. Centering-based algorithms implemented for English often use a grammatical-role based ranking, in which subject is ranked above object.

“CONSTRAINTS: For each utterance  $U_i$  in a discourse segment  $D$  consisting of utterances  $U_1, \dots, U_m$ :

1. There is at most one backward-looking center,  $Cb(U_i, D)$ .
2. Every element of the forward-looking centers list,  $Cf(U_i, D)$ , must be realized in  $U_j$ .
3. The center,  $Cb(U_i, D)$  is the highest-ranked element of  $Cf(U_{i-1}, D)$  that is realized in  $U_j$ .” (Walker *et al.* 1998: 3)

“RULES: For each utterance  $U_i$  in a discourse segment  $D$  consisting of utterances  $U_1, \dots, U_m$ :

1. If some element of  $Cf(U_{i-1}, D)$  is realized as a pronoun in  $U_i$ , then so is  $Cb(U_i, D)$ .
2. Transition states are ordered. The CONTINUE transition is preferred to the RETAIN transition, which is preferred to the SMOOTH-SHIFT transition which is preferred to the ROUGH-SHIFT transition.” (Walker *et al.* 1998: 4)

Centering transitions:

	$Cb(U_i) = Cb(U_{i-1})$ (or $Cb(U_{i-1})$ undefined)	$Cb(U_i) \neq Cb(U_{i-1})$
$Cb(U_i) = Cp(U_i)$	CONTINUE	SMOOTH-SHIFT
$Cb(U_i) \neq Cp(U_i)$	RETAIN	ROUGH-SHIFT

# Information Structure

---

*Theoretical, Typological, and Experimental  
Perspectives*

Edited by

MALTE ZIMMERMANN AND CAROLINE FÉRY

**OXFORD**  
UNIVERSITY PRESS