The influence of "aboutness" on pronominal coreference

H. Wind Cowles[1]
University of Florida, USA

Previous work examining the role of antecedent accessibility in pronominal coreference has often linked coreference to prominent structural positions that in turn are linked to information structure statuses such as topic. Three experiments examine the influence of topichood independently of structural prominence by exploring the influence of the pragmatic notion of aboutness on the written production of pronominal coreferring expressions. The results show that being mentioned in an *about*-phrase increases the likelihood that a referent will be selected as the future topic of a following sentence as well as increasing the proportion of responses with early, pronominal coreference to that referent, at the expense of coreference with the subject. These results suggest that coreference is sensitive to the status of other, structurally non-prominent referents in discourse, and that the pragmatic notion of aboutness influences pronominal coreference.

1 Introduction

Pronominal coreference poses a number of questions for language researchers, in large part because of the underspecified nature of the pronouns themselves. For example, because pronouns carry little in the way of semantic information compared to fuller forms of reference (e.g., descriptive noun phrases or names), one might expect pronouns to be more difficult to interpret and thus dispreferred as a reference form. Yet, this does not appear to be case and in fact under certain discourse conditions they appear to be strongly preferred, if not required, for coherence. This raises the question of what these discourse conditions are and what determines when a speaker or writer chooses to refer to something using a pronominal form. One potential influence comes from the information status that referents in a sentence hold: for example, a referent may be interpreted as

[1] This research was supported by start-up funds provided by the College of Liberal Arts and Sciences at the University of Florida. The author thanks Victor Ferreira for the use of his Language Production Lab for Experiment 1, and Katie Doyle (UCSD) and Nicholas Williams (UF) for their help with data collection.

what the proposition expressed by a sentence is *about* given the current discourse, in which case it is considered to be the topic of the sentence. This paper examines the influence of topic status and in particular this pragmatic notion of *aboutness* on the production of pronominal coreference.

2 Background

Many factors appear to be involved in processing pronominal coreference (see Garnham, 2001 for an overview), including the nature and position of the antecedent in the discourse (e.g., recency, frequency, grammatical role parallelism, structural prominence), the relationship between sentences containing the antecedent and the anaphor (e.g., coherence relations) and the type of predicates involved in the coreference (e.g., implicit causality). Focusing on the first type of factor, many studies have found a preference for pronominal coreference to antecedents that are mentioned prominently, either syntactically as the subject (Crawley, Stevenson & Kleinman, 1990) or in a clefted phrase (e.g., Cowles, Walenski & Kluender, in press), or linearly by being first-mentioned (e.g., Gernsbacher & Hargraves, 1988; Gernsbacher, Hargreaves & Beeman, 1989, but cf. Gordon, Hendrick & Foster, 2000; Cowles et al. in press). It is important to note that in many languages subject position is easily confounded with being first-mentioned, and in studies that have untangled these positions, there have been mixed results, with either a preference for subject antecedents (Cowles et al., in press; Kaiser & Trueswell, 2003) or indications that both factors influence pronoun interpretation independently (Jarvikivi, van Gompel, Hyona & Bertram, 2005).

The generalization made by many researchers is that pronouns are used to corefer with highly accessible antecedent referents, and these different ways of mentioning antecedents can be seen as influencing referent accessibility. A number of proposals (e.g., Ariel, 1990; Garrod, Freudenthal & Boyle, 1994; Gundel, Hedberg & Zacharski, 1993) have tied reference form to the status of antecedent referent in the cognitive representation of the discourse in this way, arguing that pronouns are used when referring to prominent or discourse-focused antecedent referents. A clear prediction of such approaches is that it is not subjecthood or first mention *per se* that is an attractor for pronominal coreference, but rather the effect that subject position or first mention has on the referent of the antecedent expression. This is in contrast to other approaches that emphasize the structural status or thematic role of the antecedent (e.g., Smyth, 1994; Chambers & Smyth, 1998). However, while the notions of cognitive status and accessibility have featured in many approaches to pronoun resolution, they are often tied to syntactic (e.g., Grosz, Joshi & Weinstein, 1995) position. If it is indeed the cognitive status of the antecedent referent that is one of the de-

termining factors in pronoun resolution, then non-syntactic manipulations of cognitive status should also influence pronoun resolution.

Another way to think about the issue of antecedent cognitive status and reference form is to consider the information structure of the utterance that contains the antecedent. For example, subjecthood is often a signal of topic status while clefts often signal (contrastive) focus status (e.g., Kiss, 1999). Word order variation, too, often has significant consequences for information structure, with fronted elements either having topic or focus status, depending on the construction used. Topic status is argued to be tied to prominent cognitive status (e.g., Lambrecht, 1999) and so is a good candidate for influencing pronominal resolution, and in fact has been included as one factor influencing pronominal coreference (Ariel, 1990). However, as we shall see below, topic status is also often associated with prominent syntactic positions, and so in order to investigate the influence of topic status while avoiding confounds with subject status or first-mention, we now turn to the pragmatic notion of *aboutness*, which can help tease apart topic status from structural prominence and primacy.

The notion of aboutness is considered to be a defining aspect of topic status by many researchers (e.g., Gundel, 1976; Reinhart, 1982, Lambrecht, 1994). In this view, the topic of a proposition expressed by an utterance is what the proposition is about, given a particular situation. That is, the topic is that part of the utterance that is the central interest or concern, and to which new information is being added (cf. Strawson, 1964). As just mentioned, topic status has been associated cross-linguistically with reduced forms of coreferring expressions, including pronouns (e.g., Ariel, 1988, 1990; Gundel et al., 1993).

Structurally, the precise mechanisms for encoding topic status differ cross-linguistically, but may involve both prosody and particular syntactic positions or constructions. In English, subject position is often associated with topic status, and may be seen as the unmarked topic position in canonical SVO sentences (Lambrecht, 1994). One classic diagnostic for topichood is the "about X" test in which a sentence is paraphrased such that the potential topic X is placed in an *about*-phrase and the felicity of this new paraphrase is determined (Gundel, 1976; Reinhart, 1982). This test is unfortunately not perfect, but suggests that mentioning a referent in an *about*-phrase may be a good way to signal (at least potential) topic status for a comprehender without using subject position, and thus it provides a way to examine topic status without confounding it with subject or other prominent syntactic positions.

Cowles (2003, Cowles & Ferreira, in prep) tested the influence of referents mentioned in *about*-phrases on the syntactic structure of spoken utterances. In one experiment, participants listened to sentences containing a target noun mentioned either in a post-verbal about-phrase (*A nurse noticed something about the lightning*.) or as the object of an embedded sentence complement (*The*

nurse noticed something as she watched the lightning.). After each sentence, they saw a theme-experiencer verb (e.g., *frightened*) followed by two nouns, one of which was the target (e.g., *baby, lightning*) and needed to verbally produced a sentence that used these three words and fit with the sentence that they just heard. Responses were coded according to whether the target noun was mentioned before (and in a higher syntactic position than) the other target noun. Crucially for present purposes, participants were instructed to use the nouns as they were presented with them, thus implicitly instructing them to not use other forms of reference to refer to the targets. Participants largely obeyed this constraint, and responses were only scored if they contained the full form of the target nouns (e.g., *lightning*). Cowles (2003) found a topic-mention advantage in which *about*-phrase targets were more likely to be produced early in an utterance compared with non *about*-phrase targets. Also, this effect was larger for theme targets than experiencer targets. These results support the idea that manipulations of topic status via *about*-phrases can have an impact on sentence production at a structural level, but do not provide evidence about whether reference form may also be affected. It is exactly this question that is addressed in the experiments that follow.

3 Experiments

3.1 Experiment 1

3.1.1 Methods

The goal of Experiment 1 was to examine whether *about*-phrases have an impact on the identification of the topic in the current sentence or influence predictions about the likely identity of the topic of the following utterance. To do this, materials were taken from Cowles (2003) and adapted for a questionnaire study in which participants were asked to indicate the topic of the current sentence and choose whether the target non-subject referent or something entirely new was most likely to be the topic of an immediately following sentence.

3.1.1.1 Participants

Forty members of the University of California, San Diego community participated.

3.1.1.2 Design and Materials

Setup sentences from experimental and filler items in Experiment 4-1 of Cowles (2003) were used. Experimental items consisted of 40 sentences that were constructed by crossing two factors: Target type (experiencer *vs*. theme) and infor-

mation status (given *vs*. about), for a total of four experimental conditions. An example is given in Table 1 below.

Table 1: Example of materials from Experiment 1

Target Type	Information Status	Setup Sentence
Experiencer	*Given*	The passenger realized something when he saw the driver.
	About	A passenger realized something about the driver.
Theme	*Given*	The passenger realized something when he saw the traffic.
	About	A passenger realized something about the traffic.

Cowles (2003) asked participants to produce sentences using theme-experiencer verbs and following this *target type* refers to whether the target argument in the setup sentence (underlined in Table 1) could be assigned the role of experiencer (e.g., driver) or theme (e.g., traffic) in the sentences that participants produced in that experiment. In all but one item target type also corresponded to a difference of animacy in which the experiencer arguments were all animate and the theme arguments were inanimate. The *about* condition was constructed by placing an indefinite noun in subject position followed by a verb followed by the word *something* and then the *about*-phase. The target argument was given as the object of the *about*-phrase. The *given* condition was constructed by using the same subject noun, but with a definite determiner, followed by a verb followed by the word *something* and then followed by a sentence complement. In this condition the target argument was always given as the object of the embedded verb.

Experimental items were divided into four lists using a Latin square design such that each item was given exactly once in each list and each list contained equal numbers of items from each condition (i.e. 10 of each condition). Forty filler items from Cowles (2003) were added to each list and then lists were pseudorandomized such that no two consecutive items were from the same condition and no more than three experimental items ever appeared in a row. Two versions of each list were created with different orders of items and fillers. This was to help prevent any spurious effects of item order within the lists.

3.1.1.3 Procedure

After giving informed consent, participants were seated in a quiet room and given the experimental materials in the form of a printed packet. All test items in the packet were presented as a sentence followed by two nouns. In experi-

mental trials, participants were given the setup sentences followed by the two possible target arguments (the theme and the experiencer). They were asked to do two things for each item in the packet. First, they needed to circle the part of the sentence that they considered to be the topic of that sentence. Then, they needed to choose one of the two following target arguments as the most likely topic of the next sentence[2]. Because of the design of the setup sentences, one of these target arguments was always previously given in the setup sentence (either in the *about* or *given* condition) and the other argument was not previously mentioned at all.

3.1.2 Results

Two measures of participant responses were calculated: the proportion of times the target argument was chosen as the topic of the setup sentence (the current topic) and the proportion of times that the *theme* argument was chosen as the most likely topic of the next sentence (the future topic). Table 2 shows the results of these measures.

[2] Participants were given the following instructions:
In linguistic theory, most sentences are considered to have "topics". The topic of a sentence is its central element, the part that any new information conveyed by the sentence is added to. For example, if you had a sentence like "As for the milkman, he noticed the yogurt had gone bad," the milkman would be the topic because he's the thing that the sentence has new information concerning, that is, that he noticed something. But, if the sentence was "As for the yogurt, the milkman noticed that it had gone bad," then the yogurt would be the topic because it is now the yogurt that is having new information added to it – the fact that it had gone bad.

In the following pages, we're going to show you the first sentence in a story. It will look something like this:

	Current Topic?	Next Topic? .
1.	As for the milkman, he noticed that some yogurt had gone bad.	o milkman o mold

We want you to do two things with each sentence: First, we want you to determine what you think the topic of each sentence is, and then circle it in the sentence. Second, imagine that another sentence is going to be written that continues the story. We want you to decide which of the two things listed next to the sentence is most likely to be the topic of that next sentence, and check the box next to your choice.

Table 2: Results from Experiment 1: Proportions of responses with the target argument selected as the current topic and proportions of responses in which the theme argument was chosen as the most likely future topic. Standard errors are given in parentheses

	Information status	
Current Topic		
Argument type	Given	About
Experiencer	.031 (.013)	.114 (.045)
Theme	.056 (.027)	.089 (.034)
Theme as Future Topic		
Argument type	Given	About
Experiencer	.292 (.039)	.181 (.038)
Theme	.647 (.045)	.817 (.039)

As Table 2 shows, for the current topic measure there was an overwhelming dispreference for the target argument, with the target circled only an average of 7% of the time (the subject of the main clause was chosen instead in all other cases). However, despite this overall dispreference, the target was still chosen relatively more often in the about condition compared to the given condition (10% vs. 4%). This effect appeared to be larger for experiencers (with an increase of 8%) than for themes (3%).

In the future topic measure there was an effect of given vs. new: participants showed a greater preference for the theme when it was previously mentioned in the setup sentence (selecting it as the most likely future topic 73% of the time) compared to when the experiencer argument had been mentioned instead (selecting the theme 24% of the time instead). There was an additional preference for the theme when it was mentioned in an *about*-phrase, which is reflected in the increase in theme selection in the theme-about condition compared to the theme-given condition. The decrease in theme selection in the experiencer-about condition compared to the experiencer-given condition also reflects a preference for selecting the about-mentioned argument as the future topic: greater experiencer selection in this condition is reflected as a decrease in theme selection.

Statistical analyses confirm these observations. For the current topic measure, two-factor (target type x information status) repeated measures ANOVAs with participants (F_1) and items (F_2) as random variables revealed a main effect of information status, such that target arguments were more likely to be chosen when they were mentioned in the about-phrase condition (F_1 (1,35) = 4.971, $p < .032$; F_2 (1,39) = 6.67, $p < .014$). There was no main effect of argument type, reflecting the fact that experiencers were no more likely to be chosen overall than themes (Fs < 1). However, there was an interaction of type and

status, reflecting the fact that the effect of status was larger for experiencers than themes (F_1 (1,35) = 4.565, p <.04; F_2 (1,39) 5.315, p < .027). Planned pair-wise comparisons show that the effect of information status was significant for experiencers (t_1(1,35) = 2.47, p < .02; t_2(1,39) = 2.78, p < .01) but only marginal for themes (t_1(1,35) = 1.78, p < .08; t_2(1,39) = 1.86, p < .07).

For the future topic measure, two-factor (target mention x information status) repeated measures ANOVAs with participants (F_1) and items (F_2) as random variables revealed a main effect of argument type (F_1(1,35) = 48.01, p < .001; F_2(1,39) = 230.6, p < .001) reflecting the preference for theme when it was previously mentioned in the setup sentence. There was no effect of information status (F_1 (1,35) = 2.36, n.s.; F_2(1,39) = 1.78, n.s.) but there was an interaction of mention and status (F_1(1,35) = 25.93, p < .001, F_2(1,39) = 36.33, p < .001), reflecting the fact that themes were chosen more when previously mentioned in the about condition and less when it was the experiencer that was mentioned in the about condition. Planned pairwise comparisons showed that the effect in information status was significant in both the theme-mentioned (t_1(1,35) = 4.29, p < .001; t_2(1,39) = 5.26, p < .001) and experiencer-mentioned (t_1(1,35) = 4.26, p < .001; t_2(1,39) = 3.67, p < .001) conditions.

The results support the hypothesis that *about*-phrases influence the topic status of their referents, with *about*-phrase referents being more likely to be chosen as topic, especially as the most likely future topic. For current topic selection, there was an overwhelming preference to select the subject of the main clause. However, aboutness still had an effect and in this case theme arguments showed a larger influence of *about*-phrase mention than experiencer arguments. Because theme arguments were inanimate in all but one item (and experiencers were always animate), this may reflect an interaction with animacy in which inanimate referents are most influenced by *about*-phrase mention, at least with respect to their information status in the current sentence. These results are similar to those reported in Cowles (2003), in which theme arguments were also more influenced by *about*-phrase mention than experiencer arguments.

3.2 Experiment 2a

Experiment 1 established that the materials from Cowles (2003) influence both current and future topic preferences, but with a much larger influence on future topic preference. Experiment 2 was designed to see whether this topic interpretation for the *about*-phrase referent would influence pronominal coreference production in a written sentence production task. Participants were given the theme setup sentences along with a theme-experiencer verb and asked to create a sentence using the verb that followed from the setup sentence.

3.2.1 Methods

3.2.1.1 Participants

Twelve members of the University of Florida community participated.

3.2.1.2 Design and Materials

This experiment had two conditions: About (when the target argument was mentioned in an *about*-phrase) vs. Given (when it was mentioned as the object in an embedded clause). The theme-given and theme-about conditions of all forty experimental items were taken from Experiment 1, as well as all filler items. Each setup sentence was paired with a theme-experiencer verb, also taken from Cowles (2003). Theme-experiencer verbs were chosen both here and in Cowles (2003) because they have been shown to be roughly equally biased in their use between passive and active voices (Altmann & Kemper, 2006; Ferreira, 1993). This was important in Cowles (2003) because the principle interest of that study was the role that topic status plays in the online production of syntactic structures, but is also useful for our present purposes because such verbs should not bias any particular argument toward subject position in the written responses elicited in these experiments.

Experimental items were divided into two lists using a latin-square design such that each item was given exactly once in each list and each list contained equal numbers of items from each condition (i.e. 10 of each condition). Forty filler items from Cowles (2003) were added to each list and then lists were pseudorandomized such that no two consecutive items were from the same condition and no more than three experimental items ever appeared in a row. Two versions of each list were created with different orders of items and fillers. This was help prevent any spurious effects of item order within the lists. Each list was formatted so that items appeared in a numbered list. Each item consisted of the setup sentence followed by the theme-experiencer verb in parentheses, presented in past tense/participle form. The verb was preceded by an arrow to help remind participants that they needed to use it in the sentence that they produced. Under each setup sentence there was a blank line for their response.

3.2.1.3 Procedure

After giving informed consent, participants were seated in a quiet room and were presented with the experimental materials in the form of a printed packet. Participants were given a set of written instructions in which they were told that they should read each sentence and following verb and then write down a sentence that used the verb and naturally followed and fit with the sentence they

had just read. They were also told that they could add new information in their sentence, but should try to keep the sentences relatively short.

3.2.2 Results

All 480 responses to experimental items were entered into a computer data file and coded in the following way. First, responses were excluded if they were ungrammatical, did not contain the given verb, or failed to use the verb appropriately (*i.e.*, as a verb). This removed 16 responses (3% of the data) from further analysis. Next, each response was coded (yes or no) for whether it contained coreference to the target noun in the setup sentence as well as whether it contained coreference to the subject. If it did contain coreference to either of these, then the form of the coreferring expression was coded (pronoun, repeated, other). Finally, each response was coded for whether the first-mentioned entity in the sentence corresponded to the subject or target in the setup sentence (or to something else). The results are given in Table 3 as proportions out of all analyzed responses[3]

Table 3: Results from Experiment 2a: Proportions of responses with coreference of subject and target arguments, both in any form as well as specifically in pronominal form, as well as proportion of first mention for subject and target. Standard errors are given in parentheses.

	Information Status of Target	
Coreference (all forms)	Given	About
Subject	.74 (.05)	.61 (.04)
Target	.45 (.04)	.60 (.04)
Coreference (pronominal)		
Subject	.69 (.07)	.56 (.06)
Target	.16 (.03)	.30 (.04)
First Mentioned		
Subject	.45 (.07)	.32 (.05)
Target	.36 (.07)	.54 (.05)

The results show that mention in an *about*-phrase has a clear effect on how the target is treated in participants' responses. First, mention in an *about*-phrase resulted in more instances of coreference in any form with the target (.60) compared to when it was merely given as part of the subordinate clause (.45). Fur-

[3] An analysis of proportion of pronominal coreference out of only those trials with coreference in any form was not possible for Experiment 2b due to items with no (esp. target) coreference. The measure of pronominal coreference out of all analyzable trials was used instead in order to allow better comparison between the experiments.

ther, *aboutness* caused an increase in use of pronouns as the form of coreferring expression. Finally, targets were also more likely to be mentioned first in the sentence when they had occurred in an *about*-phrase, in keeping with the results from Cowles (2003).

Turning to effects of target aboutness on the production of the subject, there is a similar, complementary effect. Subjects were *less* likely to be coreferenced overall when the theme had been mentioned in an *about*-phrase and were less likely to have coreference in pronominal form. Subjects were also less likely to be mentioned first when the target was previously mentioned in an *about*-phrase.

These observations are supported by statistical analyses. Paired-sample t-tests were used to compare proportions between the given and about conditions, with both participants (t_1) and items (t_2) as random factors. These analyses showed a significant difference for aboutness in all target measures: Overall coreference ($t_1(1,11) = 4.84$, $p < .001$; $t_2(1,39) = 3.36$, $p < .002$), pronoun coreference ($t_1(1,11) = 5.04$, $p < .001$; $t_2(1,39) = 4.46$, $p < .001$), and first mention ($t_1(1,11) = 4.55$, $p < .001$; $t_2(1,39) = 2.5$, $p < .02$). The analyses also showed a significant difference for overall coreference to subjects ($t_1(1,11) = 3.37$, $p < .006$; $t_2(1,39) = 3.36$, $p < .002$) as well as pronominal coreference ($t_1(1,11) = 3.2$, $p < .008$, $t_2(1,39) = 2.95$, p .007) but only a marginal difference for subject first mention ($t_1(1,11) = 1.89$, $p < .09$; $t_2(1,39) = 1.53$, n.s.).

3.3 Experiment 2b

The results of Experiment 2a suggest that aboutness influences coreference and the form of coreferential expressions. However, it is possible that the difference between the given and about conditions is not being driven by the *about*-phrase *per se* but rather by other differences between the conditions. There are two systematic differences in particular that are likely candidates: the difference in subject definiteness (subjects were always definite in the given condition and always indefinite in the about condition) and the extra coreference (always pronominal) to the subject in the given condition compared to the about condition. To examine this possibility, materials from the given condition in Experiment 2a were modified to make them more similar to the about condition and another sentence production experiment was conducted.

3.3.1 Methods

3.3.1.1 Participants

Twelve members of the University of Florida community participated. These participants did not take part in Experiment 2a.

3.3.1.2 Design and Materials

The materials were identical to those in Experiment 2a except for two changes in the given condition. First, the subject was made indefinite by using the determiner *a* instead of *the*. Second, the embedded clause was altered by removing the subject pronoun and inflecting the verb with the progressive *–ing*. Some verbs needed to be altered in order to keep the sentences as natural sounding as possible. In these cases, the verb was changed in the both the given and topic conditions. An example of the modified materials is given in Table 4 below.

Table 4: Example of materials from Experiment 2b

	Setup sentence (prompt verb)
Given	A nurse mentioned something while watching the lightning. (frightened)
About	A nurse mentioned something about the lightning. (frightened)

3.3.1.3 Procedure

The experimental procedure was identical to Experiment 2a.

3.3.2 Results

One item was dropped from analysis due to a typographical error for that item in the printed packet. All 468 responses to the remaining 39 experimental items were entered into a computer data file and coded in the same way as in Experiment 2a, with 9 responses (2%) removed from further analysis as a result of coding procedures. The results are given in Table 5.

Table 5: Results from Experiment 2b: Proportions of responses with coreference of subject and target arguments, both in any form as well as specifically in pronominal form, as well as proportion of first mention for subject and target. Standard errors are given in parentheses.

	Information Status of Target	
Coreference (all forms)	Given	About
Subject	.82 (.03)	.68 (.04)
Target	.43 (.04)	.53 (.03)
Coreference (pronominal)		
Subject	.75 (.05)	.63 (.04)
Target	.22 (.03)	.33 (.05)
First Mentioned		
Subject	.74 (.04)	.53 (.05)
Target	.18 (.03)	.31 (.04)

The results are very similar to those found in Experiment 2a: mention in an *about*-phrase had an effect on how both the target and subject were treated by participants in their response sentences. Once again, mention in an *about*-phrase resulted in more instances of coreference in any form (.53) with the target compared to when it was simply part of the subordinate clause (.43). *Aboutness* also resulted in an increase in the use of pronouns as the form of coreferring expression and targets were more likely to be mentioned first in the sentence when they had occurred in an *about*-phrase. In turn, subjects were *less* likely to be coreferenced overall when the theme had been mentioned in an *about*-phrase and were less likely to have coreference in pronominal form. Subjects were also less likely to be mentioned first when the target was previously mentioned in an *about*-phrase.

These observations are supported by statistical analyses. Paired-sample t-tests were used to compare proportions between the given and about conditions, with both participants (t_1) and items (t_2) as random factors. The difference in overall coreference to the target was marginal by participants and significant by items ($t_1(1,11) = 2.12$, $p < .06$; $t_2(1,38) = 2.10$, $p < .05$) and the same was also true for pronoun coreference to the target ($t_1(1,11) = 2.04$, $p < .07$; $t_2(1,38) = 2.62$, $p < .02$). The target was first-mentioned significantly more often when it was previously in an *about*-phrase ($t_1(1,11) = 4.52$, $p < .001$; $t_2(1,38) = 3.13$, $p < .003$). The analyses showed a significant difference for overall coreference to subjects ($t_1(1,11) = 3.23$, $p < .008$; $t_2(1,38) = 3.77$, $p < .001$) as well as pronominal coreference ($t_1(1,11) = 2.40$, $p < .04$, $t_2(1,38) = 2.65$, $p < .02$) and subject first mention ($t_1(1,11) = 4.84$, $p < .001$; $t_2(1,38) = 4.67$, $p < .001$).

4 Discussion

The results from Experiment 1 establish that the *about*-phrase manipulation had an effect on the selection of both the topic in the sentence containing the *about*-phrase and the topic of a hypothetical following sentence. While the influence on the current topic was significant, the preference for the target was very small, reflecting an overwhelming preference to view the subject of a simple, declarative sentence in a neutral context as the topic of the proposition expressed by that sentence. When presented with the choice between something previously mentioned and something new as the likely next topic of a following sentence, there was an unsurprising preference for the previously mentioned argument. However, there was an additional increase in preference for the previously mentioned argument when it had been in an *about*-phrase, suggesting that *about*-phrases do successfully influence the topic status of future mentions of the *about*-phrase referent. Additional work needs to be done to see how this preference stands up against the option of selecting the previous subject.

With the results of Experiment 1 in mind, we can turn to the written production data from Experiments 2a and 2b. The main goal in these experiments was to examine whether there would be an increase in pronominal coreference with referents that had been previously mentioned in an *about*-phrase compared to those that were not. In both experiments, the results show that there was such an increase, although it was attenuated in Experiment 2b. In fact, there was an increase in overall coreference to the *about*-phrase antecedent as well as pronominal reference. Further, an *about*-phrase antecedent was more likely to be mentioned first in the sentence than the same referent when it was not in an *about*-phrase. The pattern of results from all three of these measures is consistent with the idea that *aboutness* influences the status of referents in the discourse representation.

Interestingly, the status of the target also had an influence on coreference with the subject. In both experiments, the proportions of overall coreference to the subject as well as pronominal coreference decreased when the target was mentioned in an *about*-phrase. If the effect of *about*-phrase mention was only to increase the accessibility of the target, then one might expect to find differences only in measures of target coreference, because there is no reason why participants could not use pronouns to refer to both the subject and the target in the same response. But, clearly subject and target pronominal coreference in the responses was not independent: when target coreference increased, subject coreference decreased. This is consistent with two possibilities: first, that accessibility is related to attentional processing and thus a finite resource. From this perspective, increasing the accessibly of one referent necessarily decreases the accessibility of other referents (Arnold & Griffin, 2007) and so if aboutness increases the accessibility of a referent, then the subject referent must become less accessible. The second possibility is that *about*-phrase mention increases the likelihood that the target is interpreted as the topic of the next utterance. This topic assignment would be at the expense of the subject because there is normally only one topic in any given sentence. This second possibility, that topic status is being assigned to the target instead of the subject more often in the about condition, is supported by the fact that in Experiment 2a targets were first-mentioned more often (.54) than subjects (.32) when the target was mentioned in an *about*-phrase. Combined with the results from Experiment 1, this suggests that participants interpreted the target referent as the topic of the next sentence more often when it occurred in an *about*-phrase, and then encoded it as topic appropriately, placing it in subject position (and using pronominal coreference).

However, it is important to note that while *aboutness* had a significant effect in these different measures, *about*-phrase antecedents nearly always had lower proportions than subject antecedents. Thus, while overall it appears that *aboutness* influences coreferential form and position within a sentence, it does

not completely override the influence of being mentioned in subject position. This could be because subject position is a very strong cue to topic status in English, or it could be because of an independent structural preference for coreference with subject-mentioned antecedents that would not be diminished by a shift in topic status or accessibility of the target. That is, it may be that subject preference is really a combination of two factors: an increase in accessibility (due to topic status) and a preference for coreference with the particular structural position of subject. Further work is needed to tease these possibilities apart.

5 Conclusions

The results presented here show that the pragmatic notion of aboutness influences not only pronominal coreference to referents mentioned as objects of *about*-phrases, but also to referents mentioned first and in subject position. Although there is an overwhelming preference to interpret the subject of a sentence as the topic, *about*-phrases nonetheless appear to influence comprehenders' interpretations of both the topic of the current sentence as well as the most likely topic of a following utterance. Taken together with theoretical accounts of topichood, these results support the idea that *about*-phrases serve to guide topic interpretation, and suggest that topic status may influence pronominal coreference production even when it is not signaled via syntactic prominence or primacy. However, both of these latter factors still appear to have an important influence on the use of pronominal coreference, as the results of all the experiments showed a strong preference for using pronominal expressions to refer to subject, first-mentioned referents. Thus, while aboutness has a significant influence on pronominal coreference, it does not appear to override preferences for pronominal coreference derived from syntactic status or linear order.

6 References

Altmann, L. & Kemper, S. (2006). Effects of age, animacy and activation order on sentence production. *Language and Cognitive Processes* 21: 322-354.

Ariel, M. (1988). Referring and accessibility. *Journal of Linguistics* 24: 65-87.

Ariel, M. (1990). *Accessing noun-phrase antecedents*. London: Routledge.

Arnold, J. & Griffin, Z. (2007). The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language* 56: 521-536.

Chambers, C. & Smyth, R. (1998). Structural parallelism and discourse coherence. *Journal of Memory and Language* 39: 593–608.

Cowles, H.W. & Ferreira, V. (in prep). The influence of information structure on sentence production.

Cowles, H.W. (2003). *Processing Information Structure: Evidence from Comprehension and Production*. Unpublished PhD Dissertation, UC San Diego.

Cowles, H.W., Walenski, M. & Kluender, K. (in press). Linguistics and cognitive prominence in anaphor resolution: topic, contrastive focus and pronouns. *Topoi,* DOI 10.1007/s11245-006-9004-6.

Crawley, R.A., Stevenson, R.J. & Kleinman, D. (1990). The use of heuristic strategies in the interpretation of pronouns. *Journal of Psycholinguistic Research* 19(4): 245–264.

Ferreira, F. (1993). Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language* 33: 715-736.

Garnham, A. (2001). *Mental models and the interpretation of anaphora*. Psychology Press, Hove East Sussex.

Garrod, S. Freudenthal, D. & Boyle, E. (1994). The role of different types of anaphor in the on-line resolution of sentences in a discourse. *Journal of Memory and Language* 33: 39-68.

Gernsbacher, M.A. & Hargreaves, D. (1988). Access sentence participants: the advantage of first mention. *Journal of Memory and Language* 27: 699–717.

Gernsbacher, M.A., Hargreaves, D. & Beeman, M. (1989). Building and accessing clausal representations: the advantage of first mention versus the advantage of clause recency. *Journal of Memory and Language*, 28, 735–755.

Gordon, P.C., Hendrick, R. & Foster, K.L. (2000). Language comprehension and probe-list memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26(3): 766-775.

Grosz, B., Joshi, A. & Weinstein, S. (1995). Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics* 21(2): 203-226.

Gundel, J. (1976). The role of topic and comment in linguistic theory. Indiana University Linguistics Club, Bloomington Indiana.

Gundel, J., Hedberg, N. & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language* 69(2): 274–307.

Jarvikivi, J., van Gompel, R.P., Hyona, J. & Bertram, R. (2006). Ambiguous pronoun resolution: contrasting the first-mention and subject-preference accounts. *Psychological Science* 16(4): 260-264.

Kaiser, E. & Trueswell, J. (2003, March). The quest for a referent: Investigating the interpretation of pronouns and demonstratives in real time. Poster presented at the 16th Annual CUNY Conference on Human Sentence Processing, Boston.

Kiss, K. (1998). Identificational focus versus information focus. *Language* 74(2): 245–273.

Lambrecht, K. (1994). *Information Structure and Sentence Form: Topic, Focus and the mental representations of discourse referents*. Cambridge, U.K.: Cambridge Press.

Reinhart, T. (1982). Pragmatics and linguistics: an analysis of sentence topics. Indiana University Linguistics Club, Bloomington Indiana.

Smyth, R. (1994). Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research* 23: 197–229.

Strawson, P. (1964). Identifying reference and truth values. *Theoria* 30: 96-118.