

Discourse coherence and the interpretation of accented pronouns

Mindaugas Mozuraitis

*Cancer Care Ontario
Canada*

MOZURAITISMINDAUGAS@GMAIL.COM

Daphna Heller

*Department of Linguistics
University of Toronto
Canada*

DAPHNA.HELLER@UTORONTO.CA

Editor: Massimo Poesio

Submitted 11/2016; Accepted 08/2017; Published online 10/2017

Abstract

It has been assumed since Akmajian and Jackendoff (1970) that accenting (or stressing) a pronoun, namely making it prosodically prominent, changes its interpretation. However, more recent work suggests that not all accented pronoun receive an interpretation that is different from that of their unaccented counterpart. One proposal is that the alternative is blocked if it does not result in a plausible sentence (Taylor et al., 2013). In a series of three experiments that use an offline comprehension task, we show, first, that plausibility cannot account for the lack of reversal in all cases. We also rule out two other hypothesis about why the alternative interpretation is blocked: when it competes with a strong default, and when it is lower in salience than the default. Instead, we conclude that coherence relations indirectly constrain the interpretation of accented pronouns. Specifically, the difference between parallel and result coherence is that the former, but not the latter, explicitly provides the alternative to the accented event which drives reversal. This means that reversal is only achieved under special circumstances. Like other accented constituents, accented pronouns simply involve alternatives for their interpretation (Krifka, 2008). Our findings indicate that the most readily available alternative is the negation of the accented event; this indirectly gives rise to the effect of ‘surprise’ previously associated with accented pronouns.

Keywords: pronouns, discourse coherence, anaphora, accenting, prosody

1 Introduction

Pronouns (e.g., *she*, *he*, *it*) have received much attention in the (psycho)linguistics literature. This is because the meaning they encode is minimal (e.g., gender, number, animacy), and their interpretation depends in large part on the integration of information from the discourse context (i.e., the preceding linguistic information), and potentially other sources of information. For example, the pronoun *he* in (1b) requires a referent that is male, singular, and a person.

- (1) a. William called Oliver in the morning.
- b. Then, he went to school.

Linguistically, the pronoun *he* in (1b) is ambiguous: it can be interpreted as either of the entities mentioned in prior discourse (i.e., William and Oliver in 1a). However, there is a strong intuitive preference to interpret the pronoun as referring to William; we will call this the *default* interpretation throughout the paper. Pronoun interpretation has been accounted for using the

©2017 Mindaugas Mozuraitis, Daphna Heller

This is an open-access article distributed under the terms of a Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0/>).

concept of *accessibility* (or *salience*). Specifically, it has been widely assumed that, during the processing of discourse, listeners create a hierarchy of entities (Ariel, 1990; Gundel, Hedberg, & Zacharski, 1993; Grosz, Joshi, & Weinstein, 1995; Arnold, 2010, and many others). This hierarchy has been argued to depend on how entities were previously mentioned in the discourse, including their syntactic position, their thematic role, and also how recently they have been mentioned (Arnold, 2010). In this approach, the default interpretation of a pronoun would be the most accessible (or salient) entity in the discourse context. For example, processing (1a) would result in a hierarchy of two entities, where William is ranked higher than Oliver; this is because William was the subject of the first sentence. Next, when processing (1b), *he* will be interpreted as the higher-ranking entity in this hierarchy, namely William.

It is well known, however, that the notion of accessibility alone cannot account for the default interpretation of pronouns in all cases. For example, after processing (2a), *William* will rank higher than *Oliver* on an accessibility scale, just like after (1a), and yet the default interpretation of *him* in (2b) is *Oliver*. Similarly, (3a) would render *Bill* more accessible than *John*, but the default interpretation of *he* in (3b) is actually *John*.

- (2) a. William hit Oliver.
b. Then, Rob slapped him.
(adapted from Smyth, 1994)

- (3) a. Bill passed the comic to John.
b. Then, he read it.

(adapted from Stevenson, Knott, Oberlander, & McDonald, 2000)

These kinds of examples have led to the suggestion that pronoun interpretation should not be viewed as a process that operates on referents and their accessibility status alone; instead it should be seen as a by-product of a general mechanism of establishing coherence in discourse (Hobbs, 1978, 1979; Hobbs, Stickel, Appelt, & Martin, 1993; Stevenson, Crawley, & Kleinman, 1994; Stevenson et al., 2000; Kehler, 2002; de Hoop, 2004; Wolf, Gibson, & Desmet, 2004; Jasinskaja, Kölsch, & Mayer, 2007; Kehler, Kertz, Rohde, & Elman, 2008, among others). The idea is that the referent assigned to a pronoun is a result of a more general process of reasoning about how the meaning of the current sentence can be integrated with prior discourse. In (2), if *him* is interpreted as referring back to the previous object, this will enhance the similarity of the event in the second sentence to the event in the first sentence (cf. Hobbs, 1979; Smyth, 1994; Chambers, & Smyth, 1998; Kehler, 2002); this similarity relation between sentences has been described as a discourse relation of “parallel” coherence (e.g., Kehler, 2002). For (3), what connects the second sentence (i.e., reading a comic book) with the first sentence (i.e., giving a comic book) is world knowledge that people normally read a comic book after being given one rather than after giving one (cf. Hobbs, 1979; Stevenson et al., 2000; Kehler, 2002). This relation between sentences has been described as the discourse coherence of “result”, because of the causal relationship between the two events (e.g., Stevenson et al., 2000; Kehler, 2002).

The focus of the current paper is the interpretation of *accented pronouns* (also called *stressed pronouns*), namely pronouns that are acoustically more prominent. It has been widely accepted since Akmajian and Jackendoff (1970) that an accented pronoun receives a different interpretation than its unaccented counterpart (i.e., an interpretation different from the *default* interpretation). For example, in (2) above, the unaccented pronoun *him* is intuitively interpreted as *Oliver* (the default interpretation), while the accented counterpart *HIM* is intuitively interpreted as *William* (a *non-default* interpretation, see also Ariel, 1990; Hirschberg & Ward, 1991; Kameyama 1994, 1999; Cahn, 1995; Nakatani, 1997). This pattern of “reversal” has been accounted for with different mechanisms in the literature. For example, Kameyama’s (1999) *Complementary Preference Hypothesis* operates on a set of ‘currently salient’ entities. In this system, the interpretation of an accented pronoun is achieved by first computing the referent for the unaccented pronoun, and then substituting it with a different entity from the set of ‘currently salient’ entities. But reversal has also been derived within coherence-based accounts, such as

Kehler (2005): Here accenting does not target entities but rather events, indicating that events unfold contrary to expectations, which is achieved by assigning the accented pronoun a referent that is different from the default interpretation. Thus, despite relying on different mechanisms, the two accounts will both predict reversal in many cases that accented pronouns will receive a different interpretation than their unaccented counterparts, namely reversal.

However, not all accented pronouns actually exhibit reversal. For example, de Hoop (2004) presents a set of naturally-occurring examples from written texts, where the accented pronouns receive the *same* interpretation as their unaccented counterparts (see also German, 2009 for similar results in controlled experiments). Two of her examples are given in (4):

- (4) a. Jack and Mary are good friends. *HE* is from Louisiana.
 b. ‘I left Oxhead in the road outside. Will you see him safe for me? Take a guardor four.’
 ‘Yes, Alexander.’ He went off in a blaze of gratitude.
 There was a felt silence; Antipatros was looking oddly under his brows.
 ‘Alexander. The Queen your mother is in the theatre. Had *SHE* not better have a guard?’

In (4a), *HE* is intuitively interpreted as Jack and in (4b) *SHE* intuitively interpreted as the Queen; both are the same interpretation that would be assigned to an unaccented pronoun in the same position, or the *default* interpretation. Note that in both contexts there is only one referent that fits the gender features of the pronoun; that is, an alternative to the default interpretation is simply not available. de Hoop suggests that accenting is licensed if it can signal contrast, but this contrast need not come from a different interpretation for the pronoun itself: in (4a) Jack’s being from Louisiana is contrasted with Mary not being from Louisiana, and in (4b) the Queen having a guard is contrasted with her not having one¹. We come back to this idea in the General Discussion, but in the meantime, we may wish to draw a new generalization: the default interpretation is assigned to an accented pronoun if no alternative is available.

This idea is pursued by Taylor, Stowe, Redeker, and Hoeks (2013). In an offline judgement experiment, they examine discourses like (5a) where the coherence relation is one of parallelism, and discourses like (5b) where the coherence relation is causal, that is result.

- (5) a. Sandra called Monica, and Roger emailed her/HER.
 b. Michelle trained Beth, and Anne paid her/HER.

Taylor et al., find that in parallel discourses the accented pronoun received the non-default interpretation (i.e., a different interpretation from the unaccented pronoun), but in the result discourses there was no such reversal (cf. Tavano & Kaiser, 2008). In a separate experiment, they find that the sentence with the non-default referent was significantly more plausible in the parallel discourses (i.e., where the two sentences describe similar events) than in the result discourses (i.e., where the second event can be taken to be caused by the first). Linking the two findings, they claim that accenting a pronoun leads to a change in its interpretation only if the alternative interpretation is plausible. This proposal fits well with frameworks that subsume pronoun interpretation under the drive to make the discourse coherent (e.g., Hobbs, 1979; Kehler, 2005). However, because in Taylor et al.’s study coherence and plausibility were confounded, we cannot rule out the possibility that it is coherence and not plausibility that is responsible for the presence or absence of reversal.

The goal of the current study is to identify the reason for why some accented pronouns receive an alternative interpretation (i.e., showing reversal), while others do not. We start where

¹ We do not go into de Hoop’s analysis in more detail because it cannot account for the default interpretation of *unaccented* object pronouns in parallel discourses (such as [2] above) which are at the heart of our paper. This is because her analysis follows Centering Theory (Grozs et al., 1995) in relying on the notion of “continued topic” for pronoun interpretation, but this notion makes the wrong prediction for the parallel discourses; see Chambers and Smyth (1998) for evidence and discussion.

Taylor et al. (2013) left off, trying to dissociate discourse coherence and plausibility in Experiment 1; to this end, we use discourses that are matched on the plausibility of the two possible interpretation, but differ in coherence (parallel vs. result). Since we find reversal in parallel discourses but not in result discourses, we conduct two additional experiments that aim to explore other generalizations for the discourse situations under which reversal takes place. Experiment 2 considers the strength of the preference for the default interpretation, and Experiment 3 considers the syntactic position of antecedents. Like de Hoop (2004) and Taylor et al. (2013), we assume that an accented pronoun will receive a different interpretation if one is available, and explore the circumstances under which the alternative interpretation may not be available.

2 Experiment 1

The goal of Experiment 1 was to dissociate effects of discourse coherence from effects of plausibility. We created discourses where the final sentence contained an ambiguous (object) pronoun: consider the target sentence in Table 1. To manipulate discourse coherence, we changed the *preceding* sentence: see Table 1 (we will continue to call it *the preceding sentence* for short). To create parallel coherence, we used a preceding sentence that described an event similar to the event in the target sentence (e.g., *mailed a souvenir* and *sent a postcard*). As mentioned previously, a number of studies have shown that in parallel discourses the default interpretation of an object pronoun is the previous object (e.g., Smyth, 1994; Chambers, & Smyth, 1998). To create result coherence, we used a preceding sentence that described an event that could be taken as the cause of the event in the target sentence (e.g., Hobbs, 1979; Kehler, 2002). In these cases, the default interpretation of an *unaccented* object pronoun would be the previous subject (e.g., Wolf, Gibson, & Desmet, 2004; Kehler et al., 2008).

Introduction sentence	The animals went on a school exchange across the globe.
Preceding sentence (parallel)	Pig mailed Elephant a souvenir.
Preceding sentence (result)	Pig gave Elephant his address.
Target sentence	Then, Bear sent him/HIM a postcard.
Question	Who did Bear send a postcard?

Table 1. *Sample discourses in Experiment 1, with two possible coherence relations.*

Two aspects of the materials are worth pointing out. First, in contrast to Taylor et al. (2013), the target sentence that contains the critical pronoun is *the same* for both parallel and result coherence relations (cf. 5 above). This allowed us to keep the event described in the target sentence constant across coherence relations, which is important because some verbs could bias more towards parallel or result relations independent of the preceding linguistic context. This also allowed us to use the exact same token, and hence the same prosody, across the manipulation; this is important because some prosodic cues preceding or following the pronoun could bias towards one interpretation or another. Note that in order to keep the target sentence identical across the manipulation we used a connective that does not explicitly specify the relationship between the sentences (i.e., *then*); this meant that listeners had to *infer* the coherence relations from the events described (cf. Tavano & Kaiser, 2008). Second, because our goal was to dissociate coherence and plausibility, we aimed to create materials where, for both parallel coherence and result coherence, the two possible interpretations of the pronoun would be equally plausible – we tested the materials to verify this.

2.1 Method

2.1.1 Participants

Fifty-two undergraduate students at the University of Toronto, all native speakers of English, participated in exchange for \$5. Participants were tested individually in a session that lasted approximately half an hour. Nine additional participants were tested but were excluded from analysis because their accuracy on unambiguous comprehension questions in filler trials was below 85% (see Materials).

2.1.2 Materials (design and norming)

Sixteen three-sentence discourses were created, with some adapted from Chambers and Smyth (1998): See Table 1 for an example (the full set of items is provided in the Supplementary Material folder). The first sentence (*introduction sentence*) introduced the scenario, always referring to the animals as a group. The second sentence (*preceding sentence*) mentioned two of the three animals using their name (e.g., *Bear*, *Cat*, *Elephant*, etc.), one in subject position (*the preceding subject*) and the other in direct object position (*the preceding object*). The third and final sentence in the discourse (*target sentence*) mentioned a previously-unmentioned animal by name in subject position, and contained a linguistically-ambiguous pronoun in the object position.

There were two experimental manipulations. *Coherence* manipulated the discourse relation between the preceding sentence and the target sentence by changing the preceding sentence while keeping the target sentence identical across the manipulation. In *parallel* conditions, the preceding sentence depicted a similar event to the one depicted in the target sentence by using similar verbs (e.g., *mailed a souvenir* and *sent a postcard*). In this case, the default interpretation of an unaccented object pronoun was expected to be biased towards the preceding object (see e.g. Chambers, & Smyth, 1998). In *result* conditions, the preceding sentence described an event that could be taken as the cause of the event in the target sentence. In this case, the default interpretation of the unaccented object pronoun was expected to be biased towards the preceding subject (e.g., Kehler et al., 2008).

We verified that the parallel and result discourses did not differ with respect to the plausibility of the potential antecedents for the pronoun. In other words, our goal was to equate materials for plausibility, such that *Bear sent Pig a postcard* and *Bear sent Elephant a postcard* were both plausible continuations, for both coherence relations. To this end, we had participants recruited on Amazon's Mechanical Turk rank the discourses with a full name on a 1-7 scale, with each participant judging only one discourse – the complete details are given in Appendix A. We analyzed the data using a mixed-effects linear regression model with plausibility rating (1-7) as the dependent variable, and coherence (parallel vs. result) and antecedent (preceding subject vs. preceding object) as fixed factors (as none of the slopes significantly improved the model, so the final model only included a random intercept for items²: see more under *statistical modelling*). The model revealed that a main effect of coherence (parallel: $M = 5.88$, $SD = 1.49$ vs. result: $M = 4.87$, $SD = 1.98$, $\beta = -0.99$, $SE = 0.20$, $\chi^2 = 24.45$, $p < 0.001$), indicating that, overall, participants perceived the parallel discourses as more plausible than the result discourses. There was no main effect of antecedent ($ps > 0.5$), indicating that subject and object antecedents were equal in plausibility. Most important, the Coherence x Antecedent interaction was not significant (parallel: 5.93 vs. 5.82; result 4.95 vs. 4.80; $ps > 0.9$), indicating that the preceding subject and the preceding object were equally plausible in parallel and result discourses. In other words, this finding confirms that our parallel and result discourses did not differ in terms of the plausibility of the two potential antecedents (cf. Taylor et al., 2013).

² Because each participant judged only one story, there were no participant dependencies and hence no random effects for participants.

The *coherence* manipulation described above was crossed with the *accent* manipulation in a 2x2 within-participants design. The *accent* manipulation determined whether the pronoun in the target sentence was unaccented (or unstressed) or accented (with a narrow focus accent). Sentences were recorded in their entirety by a (female) native speaker of Canadian English who was instructed to produce a phonological distinction of accenting the pronouns (no acoustic manipulation was performed). Figure 1 plots two example sentences in the two conditions. As verification that accented pronouns indeed differed from their unaccented counterparts, we measured the duration of the pronouns and their mean pitch (F0). Indeed, the accented pronouns were longer (483 ms, $SD = 49$ vs. 287 ms, $SD = 37$) and their average pitch was higher (211 Hz, $SD = 9.9$ vs. 170 Hz, $SD = 8.8$).

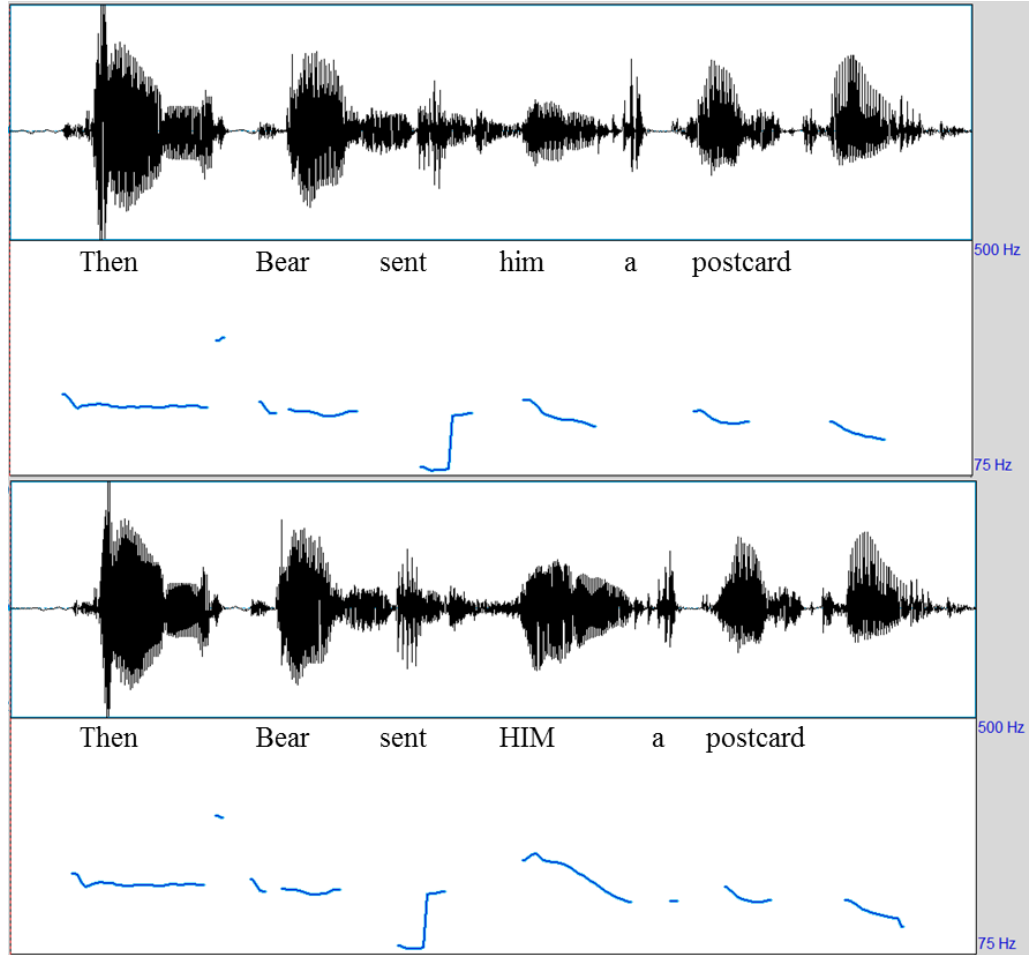


Figure 1. Sample F0 tracks for the target sentences with an unaccented pronoun (top panel) and an accented pronoun (bottom panel).

The four versions of each discourse (parallel-unaccented, parallel-accented, result-unaccented, and result-accented) were assigned to one of four presentation lists, such that each list contained an equal number of discourses in each experimental condition, and each participant encountered a specific discourse only once. Each list also included thirty-four filler discourses. The fillers were 2 to 4 sentence long. Furthermore, the filler discourses differed in structure from the experimental discourses: some contained ambiguous pronouns in subject position, some contained no pronouns at all, and some contained accented names (this was done in order to draw

attention away from accented object pronouns). The filler discourses were interspersed within the experimental discourses, such that no more than two experimental discourses occur in sequence (adjacent experimental discourses were never in the same condition).

Each of the critical and filler discourses was followed by a comprehension question. In the experimental discourses, the question targeted the interpretation of the ambiguous pronoun, and the answer was used as the dependent variable (see again Table 1). In the fillers, some questions were ambiguous whereas others were not. We used the twenty-one unambiguous questions in the filler trials as a measure that participants were paying attention; we only included in analysis those participants who answered at least eighteen questions (or more than 85%) correctly.

2.1.3 Procedure

Each participant was randomly assigned to one of the four presentation lists, with an equal number of participants in each list. Participants were told that the experiment investigated how people understood stories, and they should listen carefully so they could answer the question that would follow each story. An introductory screen displayed all six animals, e.g., “On the upper left is Elephant, to his left is Bear ...”; the pronoun *his* was used in order to ensure that all animals will be taken to be masculine. Two practice trials were used to familiarize participants with the procedure.

On each trial, pictures of the three animals appeared on the screen. After a delay of 100 ms, the discourse played over speakers. When the discourse ended, the question was displayed visually on the screen, and participants were instructed to answer it by clicking on one of the three depicted animals.

2.1.4 Statistical modelling

Given the categorical nature of the dependent variable, the data were analyzed using mixed-effects logistic regression models with participants and items as crossed, independent, random effects, implemented in package *lme4* of the statistical software R 3.2.2 (Bates, Maechler, Bolker, & Walker, 2015; R Core Team, 2015). The independent variables were coded using deviation coding: *parallel* and *unaccented* were coded as $-.5$, whereas *result* and *accented* were coded as $.5$. Pair-wise comparisons were conducted by recoding the different levels of the independent variables following West, Aiken, and Krull (1996).

We used models with the structure of random effects that was supported by the data. To select the model with the appropriate structure of random effects, we used a backwards-selection method (cf. Andrews, & Lo, 2013; Fine, & Jaeger, 2013; Taft, & Krebs-Lazendic, 2013; von Bastian, & Oberauer, 2013). Specifically, we started from a model that included the full structure of random effects supported by the design, namely random slopes for the two fixed effects and their interaction for both participants and items, and random intercepts for participants and items. We then eliminated those random effects (specifically, random slopes) that did not improve the performance of the model, starting with the interaction and following a backwards-selection procedure. At minimum, all models included random intercepts for both participants and items. Note that when conducting pairwise comparisons, the same structure of random effects was used. In the results section, we report the final structure of random effects used.

2.2 Results

The mean performance on comprehension questions in the unambiguous filler trials was 95%. The main dependent variable was how participants interpreted the ambiguous pronoun, as measured by their response to the comprehension question. We coded whether participants chose the character corresponding to the subjects of the preceding sentence (e.g., Pig in the example in Table 1) or the object of the preceding sentence (e.g., Elephant in the example in Table 1): Table 2 provides the results in terms of the mean likelihood to choose the preceding *object*. Let us first

consider the choice of antecedents when the pronoun was unaccented. In parallel discourses, the previous object was chosen 65% of the time, confirming our prediction, in accordance with previous findings in the literature, that the preferred interpretation of an object pronoun is the previous object. In result discourses, the previous object was chosen only 18% of the time; that is, the subject was the preferred antecedent, again in accordance with previous findings in the literature. Our main question, however, is what happens when the pronoun is accented. In the parallel discourses, the likelihood of choosing the preceding object was now lower (41% compared with 65%): the preferred antecedent is now the subject. In contrast, in result discourses, accenting did not seem to affect the preferred interpretation of the pronoun (20%, compared with 18%).

<i>Discourse coherence</i>	<i>Previous object choices</i>
Parallel	
Unaccented	65%
Accented	41%
Result	
Unaccented	18%
Accented	20%

Table 2. *The mean likelihood of choosing the preceding object in Experiment 1.*

Our main question is whether accenting has a different effect depending on discourse coherence, which should be reflected in an interaction between discourse relation and accenting. But because the *default* interpretation of the pronoun in the unaccented cases is different, we cannot address this question by looking at object choices directly (this could result in an interaction that arises because of the *default* interpretation, and not change due to accenting). Thus, for statistical analysis, we use a different dependent variable, noting whether participants chose the default antecedent for that context (i.e., the preferred antecedent for an unaccented pronoun). In parallel discourses, where the preceding object is the default antecedent, choosing the object is coded as 1 and choosing the subject is coded as 0 (yielding the same coding as in Table 2). In result discourses, in contrast, where it is the preceding subjects that is the default antecedent, choosing the previous object will now be coded as 0 and choosing the subject is coded as 1 (this reverses the coding from Table 2). Figure 2 plots this new dependent variable: because the analysis used is logistic regression, the data is plotted in logit space.

We fitted a mixed-effects logistic regression model (dependent variable: default interpretation = 1; non-default interpretation = 0) with coherence (parallel vs. results) and accent (unaccented vs. accented) as fixed factors. In the final model selected (see under *statistical modelling*), the random effect structure included a random slope for coherence for both participants and items, as well as random intercepts for both participants and items. The model revealed a main effect of accenting ($\beta = 0.72$, $SE = 0.18$, $z = 4.04$, $p < 0.001$), indicating that participants were overall more likely to choose the default antecedent when the pronoun was *not* accented; this is expected if accenting changes the interpretation of pronouns. There was also a main effect of coherence ($\beta = 1.66$, $SE = 0.42$, $z = 3.96$, $p < 0.001$), indicating that participants were overall more likely to choose the default referent for result discourses. Most important, these main effects were qualified by a significant Coherence \times Accent interaction ($\beta = 1.07$, $SE = 0.35$, $z = 3.02$, $p = 0.003$), indicating that accenting had a different effect on pronoun interpretation depending on discourse coherence. Specifically, accenting significantly changes the choice of antecedent in the parallel conditions (65% vs. 41%; $\beta = 1.26$, $SE = 0.23$, $z = 5.42$, $p < 0.001$), but not in the result conditions (18% vs. 20%, $p = 0.485$).

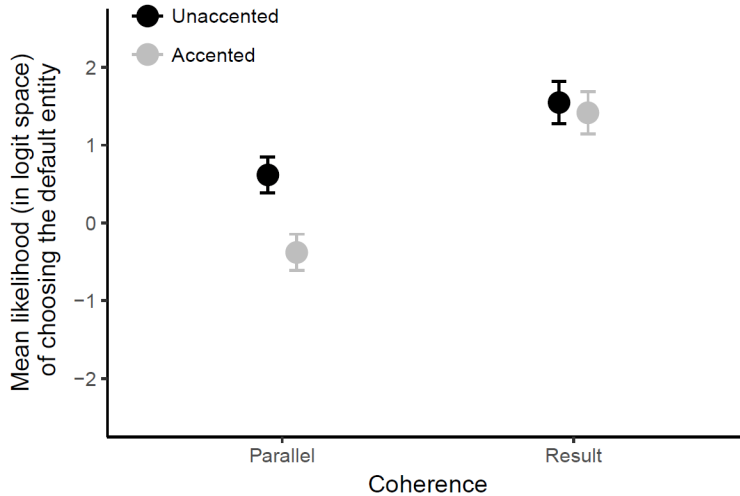


Figure 2. The mean likelihood of choosing the default entity as a function of coherence and accent, Experiment 1. This is presented in the logit space, in accordance with our logistic regression analysis (zero corresponds to 50% where there is no preference for either antecedent). Error bars represent ± 1 SE estimated for pair-wise comparisons.

We further asked whether the significant effect of accenting in the parallel condition can be characterized as a *reversal* of the preferred interpretation, namely whether the preferred antecedent changed. To address this question, we examined whether the likelihood of choosing the default antecedent in each condition differed significantly from chance; this was done using the error terms derived for the critical pair-wise comparisons above. Indeed, in the parallel-unaccented condition, participants chose the default referent significantly *more* than chance (65%, $z = 3.44$, $p < 0.001$), and in parallel-unaccented significantly *less* than chance (41%, $z = 2.37$, $p = 0.018$). These results indicate that accenting did not just have a *significant* effect on pronoun interpretation in parallel discourses, but this effect is one where the preferred antecedent is *reversed*³.

2.3 Discussion

First, these results replicate previous findings that accenting a pronoun has a different effect, depending on the coherence relations in the discourse. Like Tavano and Kaiser (2008) and Taylor et al. (2013), we find that accenting a pronoun reverses its interpretation in parallel discourses (i.e., changes which antecedent is preferred), but not in result discourses (i.e., the preferred antecedent stays the same). The current results extend previous findings by showing that (i) this effect is found even when the second event is consistent across coherence relations (cf. Taylor et al., 2013), and (ii) even when the coherence relations are not explicitly given and have to be inferred by listeners (cf. Tavano, & Kaiser, 2008). This serves as further evidence against the generalization, originally due to Akmajian and Jackendoff (1970), that what accenting does to pronouns is reverse their interpretation. As such, these results also provide evidence against theories aimed to account for a pattern of reversal, such as Kameyama's (1999) *Complementary Preference Hypothesis*.

More important, these results extend previous findings in that they allow us to assess Taylor et al.'s (2013) proposal that reversal will depend on the plausibility of the alternative interpretation. Recall that in our materials the plausibility of the alternative interpretation did not differ from the plausibility of the default interpretation, for both parallel and result discourses

³ For completeness, we note that in both result discourses, participants chose the previous subject significantly more than chance (unaccented: $z = 6.76$, $p < 0.001$; accented: $z = 6.14$, $p < 0.001$).

(even though result continuations were *overall* less plausible) – see again under *Materials*. Thus, while our findings in the main task are similar to Taylor et al.’s, they cannot be interpreted the same way: Taylor et al.’s proposal would wrongly predict reversal in our result discourses.

Having shown that plausibility is not responsible for the lack of reversal, the remainder of the paper examines other reasons for why parallel and result discourses respond differently to the accenting of pronouns. We continue to work with Taylor et al.’s logic that (i) accenting normally reverses the interpretation of a pronoun, and (i) the reason reversal is not obtained in result discourses is that the alternative interpretation is not available. What could make the alternative not available in result discourses? One possibility we consider is that the alternative interpretation is not available if there is strong preference for the default interpretation. Recall that in the unaccented condition, the preference for the default interpretation was significantly stronger in the result discourses than in the parallel discourses (82% vs. 65%; $\beta = 1.23$, $SE = 0.45$, $z = 2.48$, $p = 0.013$). It is therefore possible that the lack of reversal was because the alternative interpretation was blocked due to the default interpretation being strongly preferred. We test this hypothesis in Experiment 2.

3 Experiment 2

The goal in Experiment 2 is to examine whether the strength of preference for the default interpretation affects whether accenting creates a pattern of reversal. We test the hypothesis that a strong preference for the default interpretation renders the alternative interpretation unavailable. We focus on discourses with *parallel* relations, manipulating how much the default interpretation will be preferred – see Table 3. This is achieved by changing the introduction sentence, which either creates an expectation that the one animal has a special role (*strong preference*), or leaves it open as to which animal will participate in which role (*weak preference*). For example, in Table 3 the parallel-strong introduction sentence implies that there is one animal that went abroad, which should lead listeners to expect that actions such as *mailing a souvenir* and *sending a postcard* would be directed towards this particular character. This contrasts with the parallel-weak introduction sentence which is neutral with respect to which character may be the recipient of mail (this case is identical to the parallel discourse from Experiment 1; see again Table 1).

Introduction (parallel-weak)	The animals went on a school exchange across the globe.
Introduction (parallel-strong)	The animals were missing their friend who went abroad on a school exchange.
Preceding sentence	Pig mailed Elephant a souvenir.
Target sentence	Then, Bear sent him/HIM a postcard.
Question	Who did Bear send a postcard?

Table 3. *Sample discourses in Experiment 2 with two possible introduction sentences.*

If the alternative becomes unavailable when the default is strongly preferred, we would expect *weak* and *strong* parallel discourses to respond differently to accenting. In the strong parallel discourses, the alternative interpretation will not be available and thus an accented pronoun will receive the same interpretation as its unaccented counterpart. In the weak parallel discourses, the alternative interpretation will be available, just like in Experiment 1. If instead we find that both parallel discourses respond similarly to accenting, this will provide evidence against the preference hypothesis.

3.1 Method

3.1.1 Participants

Fifty-two undergraduate students at the University of Toronto, all native speakers of English, participated in exchange for \$5. None of these participants had participated in Experiment 1. Six additional participants were tested, but excluded from the final analyses because they did not meet criterion (85%) in answering the unambiguous comprehension questions.

3.1.2 Materials

The discourses were adapted from the parallel discourses in Experiment 1. There were two experimental manipulations. *Strength of preference* (weak vs. strong) manipulated the extent to which the default antecedent was preferred for the unaccented pronoun. This was achieved by changing the introduction sentence. To create a strong preference, the introductory sentence was changed to imply that one character is more likely to be the object of the events in the preceding and the target sentences. In the example in Table 3, it is implied that there is one animal that is away and thus should be will be receiving mail, which should lead listeners to expect that it will be the object of *mailed a souvenir* and *sent a postcard*⁴. The *weak preference* versions of the discourses were neutral in that respect (these are identical to the parallel discourses used in Experiment 1).

To ensure that our manipulation did not also alter the plausibility of the alternative interpretation, we tested the plausibility of the two continuations as in Experiment 1 by having participants from Amazon's Mechanical Turk rate the target sentence for plausibility on a scale of 1-7— see again Appendix A. Indeed, the mixed-effects regression model with plausibility rating as dependent variable and with strength of preference and antecedent as independent variables as well as random intercept for items revealed no significant main effects or interactions (p s > 0.07), indicating that the two parallel discourses, as well as both antecedents, were perceived to have the same level of plausibility.

The rest of the design of the materials for the main task was identical to Experiment 1.

3.1.3 Procedure

The procedure was identical to Experiment 1.

3.1.4 Statistical modelling

The statistical modeling was as in Experiment 1.

3.2 Results

The average performance on the comprehension questions of the unambiguous filler trials was 95%. Table 4 provides the mean likelihood of choosing the previous object. The pattern of results for unaccented pronouns suggests that our manipulation was successful: the strength of preference manipulation caused a stronger preference (weak: 65% vs. strong: 81%). This puts us in a good position to assess whether accenting is affected by the strength of the preference. We observe that when the pronoun was accented, both parallel conditions responded similarly,

⁴ For some speakers, this introduction sentence is ambiguous between the desired reading where the animals were missing one friend (i.e., the wide scope reading), and a second reading where each animal was missing a *different* friend (i.e., the narrow scope reading); consider the full list of introduction sentences in the Supplementary Material folder. Note, first, that if the narrow scope reading is available, it should create the opposite bias than the one intended: this is because it will encourage interpreting the target sentence as an event that involves a different animal (e.g., Bear sending a postcard to someone other than Elephant). The fact that our manipulation worked and the introduction sentence increased the bias as intended (81% vs. 65%) indicates that narrow scope reading is, at the very least, much less salient for most speakers.

namely with a reversal pattern (weak: 45% vs. strong: 44%). That is, the strength of the bias toward the default interpretation does *not* seem to affect the availability of the alternative interpretation.

<i>Discourse version</i>	<i>Previous object choices</i>
Weak preference	
Unaccented	65%
Accented	45%
Strong preference	
Unaccented	81%
Accented	44%

Table 4. *The mean likelihood of choosing the previous object in Experiment 2.*

For purposes of statistical analysis, we again used the dependent variable of whether participants chose the default antecedent (i.e., the preferred antecedent for an unaccented pronoun), in order to have a comparable analysis to Experiment 1 (unlike in Experiment 1, here the default antecedent was the same for both conditions). Figure 3 plots – in logit space – the mean proportions of choosing the default antecedent, or the previous object.

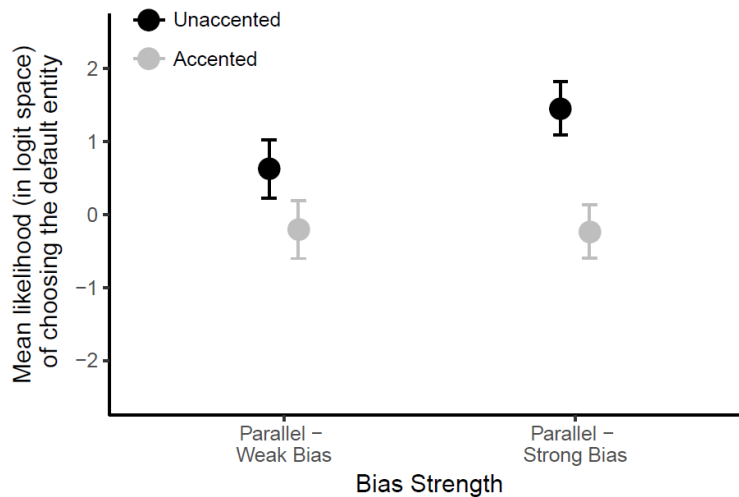


Figure 3. The mean likelihood (in logit space) of choosing the default antecedent as a function of strength of bias and accenting, Experiment 2 (chance is 0). Error bars represent ± 1 SE estimated for pair-wise comparisons.

A mixed-effects logistic regression model was fitted to data, with strength of preference (weak vs. strong) and accenting (unaccented vs. accented) as fixed factors. The random effect structure in the final model included an uncorrelated random intercept and a random accent slope for participants, and a random intercept for items (see again the model selection procedure in 2.1.4). The model revealed a main effect of strength of preference ($\beta = 0.55$, $SE = 0.19$, $z = 2.99$, $p = 0.003$), indicating that participants were overall more likely to choose the default antecedent (i.e., the previous object) in the strong preference condition than in the weak bias condition (63% vs. 55%). There was also a significant main effect of accent ($\beta = 1.84$, $SE = 0.34$, $z = 5.48$, $p < 0.001$), indicating that participants were overall more likely to choose the default antecedent

when the pronoun was unaccented (73% vs. 45%), as would be expected if accenting reversed the interpretation of pronouns. Finally, the Strength of Preference x Accent interaction was also significant ($\beta = 1.25$, $SE = 0.37$, $z = 3.34$, $p < 0.001$), indicating that the effect of accenting depended on the strength of the preference for the default interpretation. Pairwise comparisons showed that participants were significantly more likely to choose the default antecedent for the unaccented pronoun when the preference was strong, indicating that our manipulation was successful (81% vs. 65%: $\beta = 1.18$, $SE = 0.28$, $z = 4.23$, $p < 0.001$). Next, when examining the effect of accenting, we find that in both conditions an accented pronoun led participants to choose the previous object significantly *less* (strong: 81% vs. 44%, $\beta = 2.46$, $SE = 0.40$, $z = 6.17$, $p < 0.001$; weak: 65% vs. 45% $\beta = 1.21$, $SE = 0.37$, $z = 3.30$, $p < 0.001$), suggesting a pattern of reversal.

A final question is whether the effect of accenting can be characterized as *reversal*; to address this question, we ask whether the likelihood of choosing an antecedent is different from chance. When the pronoun was unaccented, the likelihood of choosing the previous object was significantly above chance in both the weak preference conditions (65%, $z = 2.44$, $p = 0.015$) and in the strong preference condition (81%, $z = 5.30$, $p < 0.001$). When the pronoun was accented, the likelihood of choosing the previous object was numerically below chance in both conditions (45% vs. 44%), but neither reached significance ($p = .395$ and $p = .285$, respectively). This means that accenting changed the preferred antecedent, but it leads to a situation of ambiguity and not strictly reversal.

3.3 Discussion

The current experiment was designed to test the hypothesis that accenting a pronoun will not change its interpretation if the alternative interpretation becomes unavailable when it is strongly dispreferred in the context. We tested this hypothesis by focusing on discourses with parallel coherence, comparing the effect of accenting across discourses that were designed to create a weak or a strong bias towards the default reading. Since the results confirm that our bias manipulation was successful (the unaccented pronoun was more likely to be interpreted as the previous object in the strong bias condition), we were able to evaluate our main hypothesis: In the accented condition, we found that accenting (numerically) reversed the preferred interpretation, independent of the strength of the bias. This pattern provides evidence against our hypothesis that a strong bias will render the alternative interpretation unavailable to be the antecedent of an accented pronoun. Instead, we observed that both discourses with parallel coherence show a (numerical) reversal pattern in response to accenting.

Having ruled out a second idea for what makes the alternative interpretation unavailable, our final step is to consider a third possible reason. We observe that what distinguishes the result discourses (Exp. 1) from the two types of parallel discourses (Exp. 1 and 2) is the syntactic position of the default antecedent: in the discourses that showed reversal, the default antecedent was the previous *object*, whereas in the discourses that did not show reversal, the default antecedent was the previous *subject*. Since it is well known that subject antecedents have a privileged status (e.g., they serve as topics and also more salient), it is possible that result discourses resist reversal because that would require switching to a *less-salient* object antecedent, whereas parallel discourses allow reversal, because here the switching is to a more salient antecedent. We test this possibility in Experiment 3.

4 Experiment 3

The goal of Experiment 3 is to examine the hypothesis that accenting would reverse the interpretation of a pronoun only if it changes from an object antecedent, which is less salient, to a subject antecedent, which is more salient, but not vice versa. To this end, we adapted the result discourses from Experiment 1 minimally by changing the preceding sentence, such that result

coherence would be best implied with the preceding *object*, not the preceding subject. If reversal was not observed for result discourses in Experiment 1 because it is not possible to switch to a less-salient alternative (i.e., an object antecedent instead of a subject antecedent), then the new result-object discourses should show reversal. This is because, like the parallel discourses, those would require switching to a *more* salient alternative, namely a subject antecedent.

4.1 Method

4.1.1 Participants

Fifty-two undergraduate students at the University of Toronto, all native speakers of English, participated in exchange for \$5. None of these participants had participated in Experiments 1 or 2. Nine additional participants were tested, but excluded from analysis, because they did not meet criterion (85%) in answering the unambiguous comprehension questions.

4.1.2 Materials

The discourses were adapted from Experiment 1: the preceding sentence from the result discourses was changed, such that the result coherence will be best implied when the antecedent of the pronoun is the previous object – see Table 5. As in Experiment 1, we tested the materials in order to confirm that all versions of discourses used in the current experiment allowed plausible readings with both antecedents (e.g., “*Then, Bear sent Pig/Elephant a postcard*”) – see the Supplementary Material folder for a list of all items. The mixed-effects regression model with plausibility rating as dependent variable and coherence and antecedent as fixed factors, as well as random intercept for items, revealed that participants perceived the result discourses as overall less plausible (parallel: $M = 5.32$, $SD = 1.82$ vs. result-object: $M = 5.88$, $SD = 1.493$; $\beta = -0.55$, $SE = 0.19$, $\chi^2 = 8.25$, $p = 0.004$). Critically, however, there was no main effect of antecedent, and the Coherence x Antecedent interaction was also not significant ($p > 0.2$). This indicates that both the parallel and the result version of the discourses were equally plausible with both possible antecedents (5.82 vs. 5.93 and 5.47 vs. 5.17, respectively). That is, the change from result-subject to result-object did not affect the plausibility of the two continuations.

The rest of the design was as in Experiment 1.

Introduction sentence	The animals went on a school exchange across the globe.
Preceding sentence (parallel)	Pig mailed Elephant a souvenir.
Preceding sentence (result-object)	Pig missed Elephant who left for India.
Target sentence	Then, Bear sent him/HIM a postcard.
Question	Who did Bear send a postcard?

Table 5. Sample discourses in Experiment 3 with two possible coherence relations.

4.1.3 Procedure

The procedure was identical to Experiments 1 and 2.

4.1.4 Statistical modelling

The statistical modeling was as in Experiments 1 and 2.

4.2 Results

The average performance on the comprehension questions of the unambiguous filler trials was 95%. Table 6 provides the mean likelihood of choosing the previous object as the referent for the linguistically-ambiguous pronoun. When the pronoun was unaccented, it was interpreted as the

preceding object 66% of the time in parallel discourses (cf. 65% for the exact same materials in both Experiment 1 and Experiment 2), and 71% of the time in the result-object discourses. This confirms that our manipulation was effective: in both the parallel discourses and the result-object discourses the default antecedent was now the previous object. As expected from previous experiments, in the parallel discourses, accenting led to reversal: the previous object was now dispreferred at 37% (cf. 41% in Experiment 1 and 45% in Experiment 2, for the exact same materials). Interestingly, the result discourses here did not exhibit the reversal effect: the previous object was still preferred at 57%.

<i>Discourse version</i>	<i>Previous object choices</i>
Parallel	
Unaccented	66%
Accented	37%
Result-object	
Unaccented	71%
Accented	57%

Table 6. *The mean likelihood of choosing the preceding object in Experiment 3.*

Here again our inferential analysis uses the dependent variable of the default antecedent, which is the previous object in both parallel and result-object discourses. Figure 4 plots the mean proportions of choosing the default antecedent for the critical pronoun, in logit space.

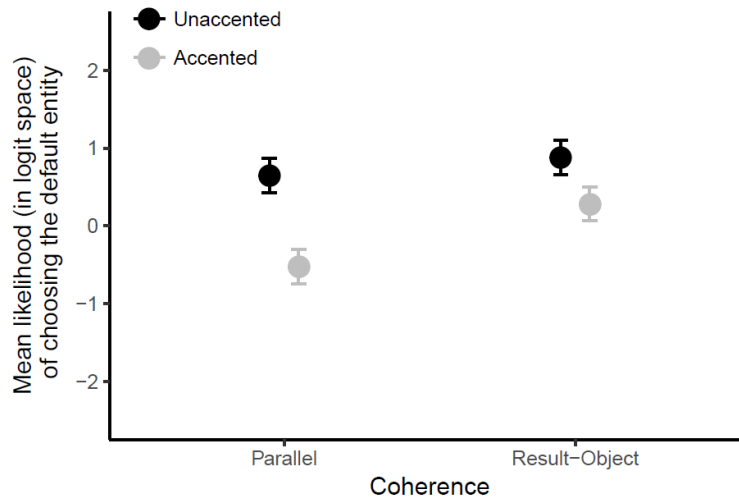


Figure 4. The mean likelihood (in logit space) of choosing the default entity as a function of accent and coherence, Experiment 3 (0 is chance). Error bars represent ± 1 SE estimated for pair-wise comparisons.

We fitted a mixed-effects logistic regression model (dependent variable: default interpretation = 1; non-default interpretation = 0) with coherence (parallel vs. result-object) and accent (unaccented vs. accented) as fixed factors. The random effect structure supported by the data was a random intercept for participants and a random intercept for items (see again the model selection process under 2.1.4). The model revealed a main effect of coherence ($\beta = 0.58$, $SE = 0.15$, $z = 3.74$, $p < 0.001$), reflecting that, overall, participants were more likely to choose the

default antecedent (i.e., previous object) in the result condition than in the parallel condition (64% vs. 51%). There was also a main effect of accent ($\beta = 0.99$, $SE = 0.16$, $z = 6.34$, $p < 0.001$), indicating that participants were overall more likely to choose the default antecedent when the pronoun was unaccented (68% vs. 47%). Importantly, these main effects were qualified by a significant Coherence x Accent interaction ($\beta = 0.66$, $SE = 0.31$, $z = 2.13$, $p = 0.033$), indicating that accenting affected pronoun interpretation differently depending on coherence relations. Specifically, although accenting had a significant effect on the choice of antecedent in both the parallel and the result-object conditions (parallel: 66% vs. 37%; $\beta = 1.32$, $SE = 0.22$, $z = 6.00$, $p < 0.001$; result-object: 71% vs. 57%; $\beta = 0.66$, $SE = 0.22$, $z = 3.02$, $p = 0.002$), the difference in parallel discourses was more pronounced than in the result discourses ($\beta = 1.38$ vs. $\beta = 0.69$).

We also tested the question of reversal, namely whether the preference is different from chance. In parallel discourses, an unaccented pronoun was interpreted as the previous object significantly more than chance (66%; $z = 2.97$, $p = .003$), and an accented pronoun significantly less than chance (37%; $z = 2.78$, $p = .005$), replicating the reversal pattern we observed in Experiment 1 (and numerically in both parallel discourses in Experiment 2). In result-object discourses, an unaccented pronoun was interpreted as the preceding object significantly more than chance (71%; $z = 4.02$, $p < .001$); the accented counterpart was also numerically above chance (57%), although this difference did not reach significance ($z = 1.43$, $p = 0.153$).

4.3 Discussion

We find that the parallel discourses exhibited a reversal pattern, but the result-object discourses do not. The fact that these discourses respond differently to accenting despite both having the preceding object as the default antecedent suggests that accenting is not simply sensitive to the syntactic position of antecedents (although it is interesting to note the contrast between a significant effect of accenting in result-object and the lack of one in result-subject).

5 General Discussion

In three experiments, we investigated why in some discourses accenting a pronoun leads to a reversal of its preferred interpretation, while in others it does not. Experiment 1 examined Taylor et al.'s (2013) proposal that reversal is blocked when the alternative interpretation is not plausible. Our results provide evidence against this proposal: when equated on the plausibility of the alternative, parallel discourses still showed reversal whereas result discourses did not. Experiments 2 and 3 investigated two other hypotheses that share the same logic, namely that reversal is not observed if the alternative interpretation is not available for some reason. Experiment 2 examined the hypothesis that the alternative is not available if the default interpretation is strongly preferred. Counter this idea, we found that a parallel discourse with a strong preference for the default interpretation nonetheless exhibited reversal. Experiment 3 examined the hypothesis that the alternative is not available if it requires switching to a less salient entity, namely from a subject antecedent to an object antecedent. Counter this idea, we found that a result discourse where the default interpretation is the preceding object resists reversal, just like result discourses where the default interpretation is the preceding subject (it is worth noting, though, that accenting had a significant effect only in the object). Note that the results of Experiment 3 provide further evidence against the hypothesis considered in Experiment 2: the absence of reversal in result-object discourses is another case where the bias is weak, but it nevertheless does not show reversal.

These results add to previous findings showing that not all cases of accented pronouns are interpreted as an alternative referent (e.g., de Hoop, 2004; Tavano, & Kaiser, 2008; German, 2009; Taylor et al., 2013). As such, they constitute further evidence that the original observation by Akmajian and Jackendoff (1970) that accented pronouns receive the reverse interpretation of their unaccented counterparts, which has been widely adopted in the literature, is an over-

generalization. We therefore argue against mechanisms that try to account for the interpretation of accented pronouns by substituting the default antecedent with a different antecedent from the list of ‘currently salient’ entities (e.g., Kameyama, 1999). Although such proposals could explain the patterns we observed for parallel discourses, they make the wrong prediction for result discourses, where accenting the pronoun did not reverse its interpretation (as well as for other cases presented in de Hoop, 2004 and German, 2009).

Furthermore, the current study also provides evidence against (three different instantiations of) a possible modification to this classical generalization: an accented pronoun will receive an alternative interpretation, unless this interpretation is blocked. The first reason for why the alternative interpretation would not be available was proposed by Taylor et al. (2013): if this reading does not lead to a plausible continuation. But the result discourses in Experiments 1 and 3 we show that a lack of reversal can be observed even when the alternative interpretation is plausible. We also proposed – and rejected – two other possible reasons. First, we considered the possibility that the alternative interpretation will not be available if it has to replace a default interpretation which is strongly preferred: The strong-bias parallel discourses in Experiment 2 exhibited (numerical) reversal despite showing a strong bias for the default, and the result-object discourses in Experiment 3 did not exhibit reversal despite only a weak bias for the default. The second reason we considered was that the alternative interpretation will not be available if it requires switching to less salient antecedent (e.g., switch from a subject to an object). The evidence against this possibility comes from the result-object discourses in Experiment 3, which did not show reversal despite the default antecedent being the object.

Let us take stock. We found reversal in two kinds of parallel discourses, repeated in (6), and no reversal in two different types of result discourse, repeated in (7).

(6) a. **Parallel-weak**

The animals went on a school exchange across the globe.
Pig mailed Elephant a souvenir.
Then, Bear sent him/HIM a postcard.

b. **Parallel-strong**

The animals were missing their friend who went abroad on a school exchange.
Pig mailed Elephant a souvenir.
Then, Bear sent him/HIM a postcard.

(7) a. **Result-subject**

The animals went on a school exchange across the globe.
Pig gave Elephant his address.
Then, Bear sent him/HIM a postcard.

b. **Result- object**

The animals went on a school exchange across the globe.
Pig missed Elephant who left for India.
Then, Bear sent him/HIM a postcard.

We thus propose that the different behavior of accented pronouns in these cases is linked to discourse coherence. But how? To answer this question, we need to consider the role of accenting beyond pronouns. It has been accepted since Rooth (1985) that the interpretation of accenting involves *alternatives* (for a review, see Krifka, 2008). While the alternative can be narrowly the antecedent alone, an accented pronoun can mark accenting on whole verb phrase (or the predicate of the sentence), and so the interpretation would involve alternatives to *sent X a postcard*⁵. We propose that parallel and result discourses differ in the alternatives that are explicitly available in the discourse context. Specifically, when alternative are called upon in order to interpret *sent HIM*

⁵ This point has been made by Kehler (2005), but he assumes that the alternative computed would involve a different antecedent for the pronoun, and thus this is essentially equivalent to computing alternatives for the antecedent alone.

a postcard, the parallel discourse contains an explicit alternative, namely the verb phrase of the preceding sentence (e.g., *mailed Elephant a souvenir*). In order for the new verb phrase to *contrast* with this alternative (cf. de Hoop, 2004), the pronoun has to receive a *different* interpretation (otherwise, the two verb phrase will be similar). In the result discourses, however, no explicit alternative is available against which *sent HIM a postcard* can be interpreted, and so listeners have to compute an appropriate alternative on the fly. Our findings suggest that the most immediate way to compute (or accommodate) an alternative is using negation: the alternative to *sent X a postcard* is NOT(*sent X a postcard*). Here the contrast is not between events with different participants, but rather between the event happening or not happening. We propose that this contrast gives rise to an inference that the event with the accenting is surprising or unexpected, a property that was proposed as part of meaning of accenting by both Kehler (2005) and German (2009).

What does this account of accented pronouns in parallel and results discourses mean for the interpretation of accented pronouns more generally? The interpretation of an accented pronoun requires contrasting the interpretation of the constituent in which the pronoun is embedded with an alternative. If an appropriate alternative is already explicitly available in the discourse context, this alternative will be used. If not, an alternative will be constructed (or accommodated), using minimal change to the current constituent, which is negation. Reversal will be observed only under special circumstances, when information in the discourse context can serve as the (contrasting) alternative with a non-default interpretation of the pronoun. It is worth noting in this context that in all experimental work on the interpretation of accented pronoun, reversal was only observed for parallel discourses (Avrutin, Lubarsky, & Greene, 1999; Venditti, Stone, Nanda, & Tepper, 2002). The other extreme is when the discourse context contains *no* information that can serve as the alternative: in this case, the alternative will be constructed (and accommodated) by applying negation to the asserted constituent. Of course, many cases are in-between: the discourse context contains information that can be readily used to construct an alternative (e.g., in de Hoop's (4a) is it that Mary is *not* from Louisiana). Our proposal is thus a departure from the standard view, originally due to Akmajian and Jackendoff (1970), that reversal is the unmarked interpretation of accented pronouns.

Acknowledgements

This work was supported by grants from the Social Sciences and Humanities Research Council to D. Heller. We are grateful to Natalie Muradian for help with stimuli preparation and testing.

Appendix A: Materials test

The goal of materials test was to confirm that the versions of discourses used in Experiments 1, 2, and 3 allowed equally plausible readings with both antecedents (e.g., Elephant and Bear) in the object position of the last sentence (e.g., “*Then, Bear sent Pig/Elephant a postcard*”).

A.1. Method

Participants. A total of 595 participants (330 male, 265 female) were recruited via Amazon’s crowdsourcing platform Mechanical Turk over the course of 30 days. Participants were paid \$0.25 each. Participants’ age was restricted to 18-50 years (mean = 31, SD = 7.45). All reported finishing high school, with 416 having attended or completed college and 79 having attended graduate school. Additional 48 participants were excluded because (i) they were slow to complete the task (3.5 SDs above the mean or 54 seconds, $n = 8$), (ii) they were extremely fast to complete the task (less than 5 seconds, $n = 35$), or because they indicated that at the age of 5 they spoke more than 2 languages ($n = 5$).

Materials and Design. The discourses used were the same as the ones used in Experiments 1, 2, and 3. Thus, there were four versions of each discourse: parallel (as in Experiments 1, 2, and 3), parallel–strong (as in Experiment 2), result–subject (as in Experiment 1), result–object (as in Experiment 3). In addition, we manipulated whether the object of the target sentence included the animal that was the preceding subject or the preceding object. In contrast to Experiments 1, 2, and 3, we used proper names and not pronouns (e.g., “*Then, Bear sent Pig/Elephant a postcard*” see Table 7 for an example). This resulted in 8 versions for each of the 16 experimental discourses.

Instructions:

In this simple task, you will read three sentences, and be asked to judge whether it makes sense for the last event (green sentence) to follow the earlier ones (blue sentences): see an example below. Read the sentences once, and choose a number on the scale that reflects **your gut feeling** about how the events fit together.

Plausibility judgment task:

Does it make sense for the last event (green sentence) to follow the earlier ones (blue sentences)?

 The animals went on a school exchange across the globe.
 Pig mailed Elephant a souvenir.
 Then, Bear sent Elephant a postcard.

 Makes no sense at all 1 2 3 4 5 6 7 Makes perfect sense

Table 7. *Example of the instructions and the plausibility judgment task.*

Procedure. The entire experiment was conducted online using Amazon’s Mechanical Turk. The whole experiment (the instructions, the sensibility judgment task, and the demographic questionnaire) was completed as a single assignment (“HIT”) that Turkers complete for payment. Before beginning the assignment, interested Turkers were presented with a description of the sensibility judgment task, an example of the task, and an electronic consent form. Table 7 provides an example of the description of the sensibility judgment task, and an example of the task. Turkers who agreed to participate were randomly assigned to 1 of the 128 discourses (4-5

participants per discourse). Each participant rated the sensibility of a single discourse to prevent against any learning and/or fatigue effects. However, across all participants, each of the 16 experimental discourses appeared in all 8 conditions approximately equal number of times (4-5 times). After completing the plausibility judgment task, participants completed the demographic questionnaire asking about age, gender, education, and language background. The entire assignment lasted approximately 3 minutes.

Each participant was exposed only to 1 passage (between-participant manipulation). Across the norming experiment, each passage used in the main experiments appeared in each of the 8 conditions (within-item manipulation).

References

- Akmajian, A., & Jackendoff, R. (1970). Coreferentiality and stress. *Linguistic Inquiry*, 1, 124–126.
- Andrews, S., & Lo, S. (2013). Is morphological priming stronger for transparent than opaque words? It depends on individual differences in spelling and vocabulary. *Journal of Memory and Language*, 68, 279–296.
- Ariel, M. (1990). *Accessing Noun-Phrase Antecedents*. London: Routledge.
- Arnold, J. E. (2010). How speakers refer: The role of accessibility. *Language and Linguistic Compass*, 4, 187–203.
- Avrutin, S., Lubarsky, S., & Greene, J. (1999). Comprehension of contrastive stress by Broca's Aphasics. *Brain and Language*, 70, 163–186.
- von Bastian, C. C., & Oberauer, K. (2013). Distinct transfer effects of training different facets of working memory capacity. *Journal of Memory and Language*, 69, 36–58.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48.
- Cahn, J. (1995). The effect of pitch accenting on pronoun referent resolution. ACL '95 Proceedings of the 33rd annual meeting on Association for Computational Linguistics. Stroudsburg, PA, USA.
- Chambers, G. C., & Smyth, R. (1998). Structural parallelism and discourse coherence: A test of Centering Theory. *Journal of Memory and Language*, 39, 593–608.
- Fine, A. B. and Jaeger, T. F. 2013. Evidence for implicit learning in syntactic comprehension. *Cognitive Science*, 37, 578–591.
- German, J. S. (2009). *Prosodic Strategies for Negotiating Reference in Discourse*. (Unpublished doctoral dissertation). Northwestern University, Evanston, Illinois.
- Grosz, B. J., Joshi, A. K., & Weinstein, S. (1995). Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 21, 203–225.
- Gundel, J., Hedberg, N., & Zacharski, R. (1993). Cognitive Status and the Form of Referring Expressions in Discourse. *Language*, 69, 274–307.
- Hirschberg, J., & Ward, G. (1991). Accent and bound anaphora. *Cognitive Linguistics*, 2, 101–121.
- Hobbs, J. R. (1978). Resolving pronoun references. *Lingua*, 44, 311–338.
- Hobbs, J. R. (1979). Coherence and coreference. *Cognitive Science*, 3, 67–90.
- Hobbs, J. R., Stickel, M. E., Appelt, D. E., & Martin, P. (1993). Interpretation as abduction. *Artificial Intelligence*, 63, 69–142.
- de Hoop, H. (2004). On the interpretation of stressed pronouns. In R. Blutner and H. Zeevat (eds.), *Optimality Theory and Pragmatics* (pp. 25–41). Palgrave Macmillan. New York.
- Jasinskaja, E., Kölsch, U., & Mayer, J. (2007). Nuclear accent placement and other prosodic parameters as cues to pronoun resolution. In Branco, A. (ed.) *Anaphora: Analysis, Algorithms and Applications* (pp. 1–14). Berlin/Heidelberg: Springer.
- Kameyama, M. (1994) Stressed and unstressed pronouns: Complementary preferences. In Bosch,

- P. and van der Sandt, R. (eds.), *Focus and Natural Language Processing* (pp. 475–484). Institute for Logic and Linguistics, IBM, Heidelberg.
- Kameyama, M. (1999). ‘Stressed and unstressed pronouns: complementary preferences’. In P. Bosch and R. van der Sandt (eds.), *Focus: Linguistic, Cognitive, and Computational Perspectives* (pp. 306–321). Cambridge: Cambridge University Press.
- Kehler, A. (2002). *Coherence, reference, and the theory of grammar*. Stanford, CA: CSLI Publications.
- Kehler, A. (2005). Coherence-driven constraints on the placement of accent. In E. Georgala and J. Howell (eds.), *Proceedings of the 15th Conference on Semantics and Linguistic Theory (SALT-15)*. CLC Publications. Cornell University. 98–115.
- Kehler, A., Kertz, L., Rohde, H., & Elman, J. L. (2008). Coherence and Coreference Revisited. *Journal of Semantics*, 25, 1–44.
- Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica*, 55(3), 243–276.
- Nakatani, C. H. (1997). *The computational processing of intonational prominence: A functional prosody perspective*, PhD thesis, Harvard University, Cambridge, Massachusetts.
- R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rooth, M. (1985) *Association with Focus*, PhD thesis, GLSA, University of Massachusetts, Amherst.
- Smyth, R. (1994). Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research*, 23, 197–229.
- Stevenson, R. J., Crawley, R. A., & Kleinman, D. (1994). Thematic roles, focus, and the representation of events. *Language and Cognitive Processes*, 9, 519–48.
- Stevenson, R. J., Knott, A., Oberlander, J., & McDonald, S. (2000). Interpreting pronouns and connectives: Interactions among focusing, thematic roles, and coherence relations. *Language and Cognitive Processes*, 15, 225–262.
- Taft, M., & Krebs-Lazendic, L. (2013). The role of orthographic syllable structure in assigning letters to their position in visual word recognition. *Journal of Memory and Language*, 68, 85–97.
- Tavano, E., & Kaiser, E. (2008). *Effects of stress and coherence on pronoun interpretation*. Poster presented at the 21st annual CUNY Conference on Human Sentence Processing, Chapel Hill, NC.
- Taylor, R. C., Stowe, L. A., Redeker, G., & Hoeks, J. C. J. (2013). Comprehension of marked pronouns in Spanish and English: Object anaphors cross-linguistically. *The Quarterly Journal of Experimental Psychology*, DOI:10.1080/17470218.2013.773356
- Venditti, J. J., Stone, M., Nanda, P., & Tepper, P. (2002). Discourse constraints on the interpretation of nuclear-accented pronouns. In *Proceedings of the 2002 International Conference on Speech Prosody*. Aix-en-Provence, France.
- West, S. G., Aiken, L. S., & Krull, J. L. (1996). Experimental personality designs: Analyzing categorical by continuous variable interactions. *Journal of Personality*, 64, 1–48.
- Wolf, F., Gibson, E., & Desmet, T. (2004). Discourse coherence and pronoun resolution. *Language and Cognitive Processes*, 19, 665–675.