Contrastive prosody and the subsequent mention of alternatives during discourse processing

Amy J. Schafer, Amber Camp, Hannah Rohde, and Theres Grüter

A. Schafer: Department of Linguistics, University of Hawai‘i at Mānoa, USA

amy.schafer@hawaii.edu (808) 956-3226

A. Camp: Department of Linguistics, University of Hawai‘i at Mānoa, USA acamp@hawaii.edu

H. Rohde: Linguistics and English Language, University of Edinburgh hannah.rohde@ed.ac.uk

T. Grüter: Department of Second Language Studies, University of Hawai‘i at Mānoa, USA

theres@hawaii.edu

Author’s final copy (2018-05-31). To appear in: K. Carlson, C. Clifton Jr., & J. Fodor (Eds.) Grammatical Approaches to Language Processing – Essays in Honor of Lyn Frazier. Springer: Studies in Theoretical Linguistics.

Abstract

Linguistic research has long viewed prosody as an important indicator of information structure in intonationally rich languages like English. Correspondingly, numerous psycholinguistic studies have shown significant effects of prosody, particularly with respect to the immediate processing of a prosodically prominent phrase. Although co-reference resolution is known to be influenced by information structure, it has been less clear whether prosodic prominence can affect decisions about next mention in a discourse, and if so, how. We present results from an open-ended story continuation task, conducted as part of a series of experiments that examine how prosody influences the anticipation and resolution of co-reference. Overall results from the project suggest that prosodic prominence can increase *or* decrease reference to a saliently pitch-accented phrase, depending on additional circumstances of the referential decision. We argue that an adequate account of prosody's role in co-reference requires consideration of how the processing system interfaces with multiple levels of linguistic representation.

1 Introduction

A long line of influential research, including a substantial number of works by Frazier and her collaborators and students, has established the importance of prosody and intonation in sentence and discourse processing (see Carlson, 2009 for a review). Among other findings, these studies have shown that prosodically prominent phrases attract modification and attention (Fraundorf, Watson, & Benjamin, 2010; Schafer, Carter, Clifton, & Frazier, 1996), facilitate the processing of alternatives (Dahan, Tanenhaus, & Chambers, 2002; Husband & Ferreira, 2016; Ito & Speer, 2008) and support the postulation of parallel information structure in constructions with ellipsis (Carlson 2002; Carlson, Dickey, Frazier, & Clifton, 2009; Carlson & Harris, 2018).

Evidence has been more elusive for an effect of prosody on the processing of co-reference, with studies of next mention often showing only weak or partial effects of prosodic manipulations (Balogh, 2003; Cowles, Walenski, & Kluender, 2007; Itzhak & Baum, 2015; Kaiser, 2010, 2011). This is surprising considering the strong relationship between prosody and information structure and between non-prosodic information structure and co-reference (Arnold, 2010; Colonna, Schimke, & Hemforth, 2015; Rohde & Kehler, 2014; Schumacher, Backhaus, & Dangl, 2015). However, recent experimentation in our laboratory has revealed clear effects of prosodic prominence on next-mention preferences, in native speakers of English and in second-language learners, a population often found to have difficulty with prosody in the target language (Schafer, Takeda, Camp, Rohde, & Grüter, 2015; Schafer, Takeda, Rohde, & Grüter, 2015). In the 2015 studies we manipulated the placement of prosodic prominence between two different positions in a critical sentence and found – along the same lines as the earlier studies, but more robustly – that subsequent reference echoed the location of prosodic prominence.

While these findings might seem to have established that accentuation facilitates co-reference, at least under well-controlled conditions, an apparent conflict remains between this pattern of results and some of the claims in the linguistic literature, summarized below, which lead to a prediction for different patterns of next mention. Here, we explore the relationship between prosodic prominence and next mention by considering two different explanations for how prosody might influence co-reference. We then present results from a new experiment that uses the same critical recordings as the 2015 studies, but shifts the discourse environment, and reveals a different pattern of next mention preferences. More specifically, the new results show that prosodic prominence can also facilitate next mention of alternatives to the prominent argument, which co-occurs with *reduced* co-reference to the prominent material. We interpret these findings as consistent with the view that prosodic distinctions are interpreted at many levels of linguistic analysis, and in light of the specific discourse situation in which they occur.

1.1 Salience and Alternatives

Consider a discourse fragment such as (1), in which the capitalization of *LAURA* indicates that it carries the most perceptually prominent pitch accent in the sentence. Such a pronunciation will make the phrase *Laura* more acoustically salient than it would be in a more neutral delivery. In English it will typically be longer in duration, have greater amplitude, and have a higher intonational peak (Breen, Fedorenko, Wagner, & Gibson, 2010). This salience could also be indicated at other levels of linguistic representation. For example, the presence of a pitch accent on *Laura* allows it to be represented as being part of the focused material at the syntactic and semantic levels (Selkirk, 1984). Analyses of discourse processing postulate that prominent material can be readily chosen as the topic of the next sentence in the discourse (e.g., is marked as

a *forward-looking center*, Gordon, Grosz, & Gilliom, 1993), often using a gradient treatment of prominence in which many factors can contribute to what is variously characterized as prominence, salience, accessibility, or the focus of attention (Arnold, Kaiser, Kahn, & Kim, 2013). At the point that a comprehender encounters *She* in (1), the relatively high salience at one or more levels of representation of the phrase *Laura* or the entity denoted by *Laura* could increase the likelihood of it being chosen as the antecedent. Thus, under this hypothesis, which we will call the *Salience hypothesis*, prosodic prominence leads to increased reference to the accented phrase – exactly as we saw in our earlier results.

(1) Sue threw LAURA a purple hat. She…

Yet focus, as a semantic construct, is not the mere presence of salience. Semantic focus can be analyzed as an implicit selection of the focused element from a set of contextually plausible alternatives (Roberts, 2012; Rooth, 1992). Thus, in (1), the entity denoted by *Laura* is set against other plausible recipients of the hat (perhaps Chuck or Janet), even if there is no explicit mention of these alternatives in the discourse. This type of alternative-set analysis is supported by evidence from psycholinguistic experiments showing that prosodic prominence facilitates the activation of alternatives to the accented element (Braun & Tagliapietra, 2011; Husband & Ferreira, 2016; Ito & Speer, 2008) as does the presence of a focus operator (Kim, Gunlogson, Tanenhaus, & Runner, 2015).

More generally, Roberts (2012) argues that focus supplies an answer from a presupposed set of alternatives to an implicit Question Under Discussion (QUD), and prosodic prominence helps define the set of alternatives. Describing patterns that have long been discussed in the

semantic literature (e.g., Jackendoff, 1972; Steedman, 2014), Roberts further argues that QUDs can be structured to include super-questions, such as *Who threw what to whom?* and sub-questions, such as the set: {*What about Laura – what did Sue throw HER?, And what about Jane – what did Sue throw HER?, …*}, built around the alternative set established by the prosodic prominence in (1). Because a discourse fragment like (1) can establish a set of sub-questions, it is possible to construct continuations that supply the answer to a second sub-question, for example by continuing: *She… threw JANE a bright pink skirt*. In short, the activation of alternatives evoked by semantic focus can lead to continuations that mention some of these alternatives, which we will refer to as the *Enumerated Alternatives hypothesis*. Under this hypothesis, prosodic prominence can lead to the mention of an entity that contrasts with the original phrase, and therefore reduced reference to the accented phrase – the opposite of what is predicted by the Salience hypothesis.

Interpretations that involve such sub-questions of the QUD are particularly associated with the very contours that our 2015 studies employed to instantiate prosodic prominence: L+H* L-H% tunes (following the ToBI system; Beckman & Ayers Elam, 1997), also referred to as a rise-fall-rise contour (Constant, 2012) or Jackendoff's Accent B (Jackendoff, 1972). L+H* L-H% contours are realized in these cases on what are known as contrastive topics, but this tune can be used more generally when there is a combination of focus and continuation (Dennison, 2010; Dennison & Schafer, 2017; Kurumada, Brown, Bibyk, Pontillo, & Tanenhaus, 2014; Pierrehumbert & Hirschberg, 1990).[1]

[1] Contrastive topics, Accent B, and the rise-fall-rise contour are also instantiated with the L*+H L-H% contour. The prosody-meaning relationship is complicated by the fact that many intonational descriptions in the literature are impressionistic and the inventory of pitch accents in English has been subject to debate (e.g., Calhoun, 2010). See Dahan (2015) for a review of the relationship between prosody and information structure, and the connection of prosody to constructs such as the theme/rheme distinction versus the alternative-semantics notion of contrast or 'kontrast'.

Despite the long-observed connection between our critical tune and interpretations involving sub-questions, we saw no evidence for Enumerated Alternatives in the Schafer, Takeda, Camp, et al. (2015) results (or the related findings from non-native speakers). Several aspects of that work may have discouraged such responses, though – a point we return to below. Critically, the experiment reported in this chapter was designed to provide ample opportunity for continuations that instantiated alternative sub-questions of the QUD, while using the same L+H* L-H% realizations of prominence as the earlier experiment, via truncated copies of the same soundfiles. We tested whether we would once again see evidence for the Salience hypothesis, or would instead find support for Enumerated Alternatives. Put differently, we examined how listeners would resolve the interpretation of acoustic prominence, realized with a pitch accent strongly associated with contrast, in a discourse context that made available an explicit set of alternatives: would listeners continue to take the tune as indicating general prominence or salience of the accented material, or would they now perceive the tune as establishing contrastive alternatives and inviting subsequent mention of them?

2 Experiment

We tested the Salience hypothesis and Enumerated Alternatives hypothesis with an open-ended story continuation task (cf. Stevenson, Crawley, & Kleinman, 1994), in which participants received the beginning of a story (*Sue threw Laura…*) and then added additional material to continue the story, using whatever form of expression they preferred. Unlike previous implementations of this task, and our own use of it in the 2015 studies, we implemented several features to promote the availability of contrastive alternatives, in order to test the interpretation of prosodic prominence in a richer referential context. First, we began each trial with a

background sentence that provided an explicit set of alternative characters (see (2)). Second, we displayed these characters and labels showing their names (see Fig. 1) throughout the time the participants created their continuations. We reasoned that both of these steps would allow easy access to and mention of alternative characters. Third, the displays included two non-human entities (a hat, a skirt; Fig. 1) that represented a second alternative set for the critical sentence, readily allowing the type of paired foci (*Sue threw Laura a purple hat, and she threw Jane a bright pink skirt*) found with the Enumerated Alternatives interpretation. Fourth, we presented incomplete sentences (*Sue threw Laura…*) and asked participants to finish each sentence with one of the paired non-human options before creating the independent portion of their continuation. This accomplished two goals. It forced the participants to consider the second alternative set, and it removed the sentence-final fall (L-L%) of the 2015 stimuli, which could have discouraged the postulation of alternative events (Dennison, 2010); a phrase-final rise is often used in lists and other situations in which the speaker wishes to indicate incompleteness and connection to subsequent material (Pierrehumbert & Hirschberg, 1990). And finally, the freedom to produce continuations of any form allowed participants to use a contrastive connector or mention an alternative character, to more easily supply an answer to an alternative sub-question.

We assumed that this combination of changes would greatly increase the availability of Enumerated Alternative interpretations, and thus provide a measure of the relative preference of these interpretations versus ones in which the story continues with non-contrastive discourse relations – e.g., continuations that describe a simple result or outcome or an explanation for the event (Kehler, 2002), all of which are compatible with the type of next-mention choices predicted by the Salience hypothesis, and were common in our earlier results.

To our knowledge, this type of configuration had never been tested. If a set of alternative sub-questions is strongly evoked by the critical tune, we expected to find frequent mention of contrastive alternatives in our participants' continuations. We speculated that this should be especially true if a set of alternative sub-questions is treated by the processing system as a type of relation that must be completed, e.g. if rapid mention of each relevant sub-question would ease the memory burden associated with an incomplete relationship (cf. Frazier, 1987).

2.1 Participants

The experiment was conducted with 43 members of the University of Hawai‘i community, all of whom identified as native speakers of English, gave informed consent, and were compensated with a small amount of course credit. Of these, four participants were removed from the analysis for failure to follow the task instructions and one because of equipment failure, leaving 38 participants in the final analysis.

2.2 Materials and Design

Materials for five practice, 18 critical, and 40 filler stimulus sets were constructed from previous recordings. The stimuli for each trial consisted of a paired audio recording (see (2)) and visual display (see Fig. 1) that presented a narrative event, initiated in the audio recording and continued by the participant to create a short discourse. The audio materials began with a background sentence (2a) that mentioned three human characters by first name in a conjoined noun phrase (NP). On critical trials, these conjoined NPs were never the initial phrase of the background sentence and never served as its syntactic subject, to avoid placing any of the names in an especially prominent position. A second sentence, which we will refer to as the context

sentence, continued the story but was truncated so that participants would have to complete it to carry on the story.

Critical context sentences (2b) repeated two of the established names, one in the syntactic subject position and a second in the indirect object position of a double-object transfer-of-possession sentence. These arguments served respectively as the Source and Goal of the transfer event. The order in which the three names were mentioned in the background sentence versus the context sentence was balanced across items, so that each of the six possible name orders was utilized with three critical items. Critical items used one of nine common transfer-of-possession verbs, selected to be familiar to learners of English tested in related experiments. Each verb was used twice, with different arguments and background sentences. Critical sentences were truncated at the offset of the indirect object, and so prior to mention of the Theme of the transfer event. Practice and filler sentences used a range of syntactic forms and were truncated just prior to the last NP of the sentence.

(2) Sample item (a: background sentence; b: truncated context sentence)

a. The school had asked Sue, Jane, and Laura to create the children's costumes for the holiday performance.

b. $\text{Sue}_{\text{Source}}$ threw $\text{Laura}_{\text{Goal}}$ (…)

Fig. 1 Sample visual display

Visual stimuli depicted five entities: the three human characters mentioned in the background sentence, and two objects or locations that were plausibly associated with the narrative event. In each visual scene, the three human characters were located in a left-to-right configuration that matched their order of mention in the background sentence. For the context sentence associated with the scene in Fig. 1, left-to-right, Sue is the Source of the transfer event, Jane is an unmentioned person, which we call the Alternative Person, and Laura is the Goal. The remaining two entities were placed in the bottom row. On critical trials, these two entities depicted two plausible Themes for the critical sentence. On filler trials, they depicted plausible objects or locations for the described event. Each depicted entity was labeled with a unique first name (e.g., *Sue*) or short description (e.g., *a purple hat*).

The 18 critical trials were realized in one of three conditions, which varied whether the original recording of the truncated context sentence presented Broad prominence, Source prominence, or Goal prominence, described further below. These three conditions were distributed across three presentation lists in a Latin square design. Each presentation list employed the same pseudorandomized order of critical and filler trials, constrained so that at least one filler occurred between each pair of critical trials and so that the two instances of each critical context verb were divided between the first and second halves of the experiment.

All auditory stimuli had been produced by a native speaker of American English, using a clear speaking style appropriate for non-native listeners and also deemed natural for native listeners.[2] Critical context sentences were pronounced with one of three patterns of prosodic

[2] Sample recordings are available at http://www2.hawaii.edu/~aschafer/snds.html#GRS.

prominence. In Broad prominence context sentences, each content word received a light pitch accent, but no word was uttered with contrastive prosody. In Source prominence sentences the Source NP carried a prominent L+H* pitch accent followed by an L-H% rise. A similar contour was placed on the Goal in Goal prominence conditions. The L+H* pitch accent is commonly described as inviting contrastive focus, and the L-H% sequence suggests some kind of incompleteness (Pierrehumbert & Hirschberg, 1990). As noted above, L+H* L-H% tunes are a type of rise-fall-rise contour associated with (but not limited to) contrastive topics, which have been analyzed as including a nested marking of focus (Constant, 2012; Roberts, 2012; Tomioka, 2010). The Source prominence and Goal prominence conditions will be referred to collectively as the contrastive conditions, versus the non-contrastive Broad prominence condition. Each token in a contrastive condition provided a salient indication of prosodic contrast on the prosodically prominent phrase, along with less prominent pitch accents on each content word of the surrounding material, in keeping with the information structure of the discourse. Further information about the intonational transcription and acoustic analyses of the stimuli can be found in Schafer, Takeda, Camp, et al. (2015).

2.3 Procedure

The experiment took place in a sound-attenuated booth equipped with a computer running E-Prime 2.0 software (Psychology Software Tools, Pittsburg, PA), desktop speakers, and a lavalier microphone. Participants were told that they would hear a series of incomplete stories, accompanied by visual displays, and that some of the entities depicted in the displays would be mentioned in the stories, and some would not. Each trial began with the simultaneous presentation of a visual scene and the audio recording of the initial portion of a story, played at a

comfortable volume over the speakers. The recordings consisted of the background sentence, a 750 ms silent interval, and the incomplete context sentence. At the offset of the audio recording, the participant orally completed the second sentence by naming aloud an object or location from the labelled options, with the supplied label, and then continued the story with whatever came to mind. The experimental software automatically created audio recordings of up to 30 seconds for each trial, capturing the playback of the audio stimuli and the subsequent continuation by the participant. The audio playback for critical trials averaged less than 8s, leaving ample time for the participant to produce a continuation during the recorded interval. The trial and the recording terminated when the participant pressed the space bar, which initiated a 500-ms inter-trial interval and advanced the experiment to the next trial. Participants were instructed to avoid adding humor and to treat each trial as a separate story from the others. Experimental sessions took about 50 minutes to complete, including the consent and debriefing processes.

2.4 Data Transcription and Annotation

Continuations for the critical items were first transcribed into standard English orthography and then annotated for linguistic properties. Twenty-three trials were eliminated for errors (critical portions were inaudible, the response was incoherent or confused the names, or the participant failed to supply a continuation). Recall that participants were left free to choose the syntactic form for their continuations, so that the data would be unconstrained by biases introduced by factors such as a requirement to begin a new sentence or to do so by using a subject pronoun. The continuations thus took a wide variety of syntactic forms, as illustrated by the data in (3), which are a sample of continuations for (2).

(3) Sample continuations for *Sue*$_{Source}$ *threw Laura*$_{Goal}$ *(…)*

a. …a bright pink skirt. Sue loves the color pink.

b. …a bright pink skirt. Laura threw Sue a purple hat.

c. …a purple hat while Jane suggested that a bright pink skirt would go great with it.

d. …a purple hat that they thought might be appropriate for the married character in a holiday play.

e. …a purple hat because the main idea of the play was *Alice in Wonderland* and the actors needed to be properly outfitted in English attire.

f. …a bright pink skirt and told her to put it on while she put on a purple hat.

To keep annotation and analysis well-defined across this richness of form, we focused on two simple measures. First, to evaluate preferences for next mention of alternatives to the prominent Source or Goal, we annotated the referent of the first explicit reference to a human following the selected Theme (recall that two plausible themes were provided in the visual display, each with a written label). We will refer to this measure as *First Mention*. Data were annotated for whether the First Mention was to one of the three provided characters (the Source: 3a, the Goal: 3b,f, or the Alternative Person: 3c) or to a group of them (Plural: 3d), or if there was no reference to a person in the recorded portion of the continuation or ambiguous reference (Other: 3e).

Second, to evaluate whether the continuations could be taken as addressing sub-questions of the QUD, we examined whether they contained any type of potentially parallel event to the context sentence with a contrasting entity. In determining parallel events, we included continuations that described another transfer-of-possession event and also any that could be seen

as a series of alternative sub-events in the situation, even if they did not involve transfer. For example, the continuation in (3b) presents a fully parallel transfer event with alternative entities in each argument position. The one in (3c) describes another sub-event in the task of creating costumes, which mentions Jane in contrast to Sue and the skirt in contrast to the hat. Although our detailed coding examined contrast by syntactic position, many of the continuations provided some kind of contrast but used a non-parallel syntactic form, as in (3c) or examples such as *Kevin presented CRAIG a special award, and then MICHAEL received a huge trophy*, which places an alternative Goal in the syntactic subject position. To prevent a proliferation of statistical analyses, adequately capture contrastiveness across such non-parallelism, and provide a liberal count of alternative sub-questions, we collapsed the sub-coding of contrasting entities across syntactic position into a general binary category of whether the continuation expressed *Contrast* or not.[3]

All continuations were annotated by one researcher, blind to the prominence condition. Three additional researchers, also blind to condition, each annotated a separate random 10% of the data. Agreement was over 90% for each pair of annotators for each annotation category.[4]

2.5 Results

Previous research using transfer-of-possession sentences in story continuation tasks has established a preference for the next mention to refer to the Goal, at least when the continuation

[3] Another common dependent measure for story continuation tasks is the form of the referential expression (e.g., pronoun versus name). Although we did annotate the data for referential form choice, the open-ended nature of the continuation introduces additional influences on the form of the first mentioned person, such as the distance from its antecedent and shifts in syntactic position, which would take the discussion beyond the central research question of this chapter. Likewise, it was infeasible to sub-divide the data with respect to whether the continuation began a new sentence or discourse unit (Colonna et al., 2015).

[4] Cohen's kappa scores for the first versus second annotation were .983 for First Mention and .764 for Contrast (without applying any correction for the prevalence of no-contrast responses; see Table 2), indicating acceptably high agreement.

begins a new sentence (Arnold, 2001; Grüter, Rohde, & Schafer, 2017; Stevenson et al., 1994), and particularly when the event is described as completed, as in the materials used here (Kehler, Kertz, Rohde, & Elman, 2008). We expected to find a similar general preference for Goal mention in the present experiment, which was indeed borne out. Responses, summarized in Table 1, showed a mix of choices across the annotation categories but Goal mentions nevertheless dominated the responses in each prominence condition, with averages two to three times the proportions for each of the remaining categories.

Table 1 Mean proportions of First Mention by prosodic prominence

	First Mention				
Prominence	Goal	Source	Alt. Person	Plural	Other
Broad	0.44	0.14	0.13	0.15	0.14
Goal	0.35	0.16	0.17	0.14	0.18
Source	0.38	0.16	0.14	0.16	0.16

Since Goal choices for First Mention were both strongly expected and provided the closest proportions to .5 in the obtained data, we analyzed the data in a mixed-effects logistic regression model with maximal random effects that compared Goal First Mention to any other response, using lme4 (Bates, Maechler, Bolker, & Walker, 2015) in R (R Core Team, 2017). The prominence factor was given simple coding, so that the model intercept would reflect the grand mean and each contrastive condition could be compared to the Broad prominence condition. Under the Salience hypothesis, Goal First Mentions should be higher with Goal prominence than with Broad prominence, and lower with Source prominence than with Broad prominence, since

prominence draws reference. According the Enumerated Alternatives hypothesis, the patterns should be in the opposite direction, because prominence leads to the listing of alternatives with contrasting arguments.

The results showed that Goal First Mention was significantly lower with Goal prominence than with Broad prominence (Goal prominence: .35 vs. Broad: .44, $\hat{\beta}$=-0.47, z=-1.99, p<.05), which is in line with the Enumerated Alternatives hypothesis and in opposition to the Salience hypothesis. Numerically, the Source prominence condition resulted in less frequent Goal First Mention than the Broad Prominence condition (Source prominence: .38 vs. Broad: .44), as predicted by the Salience hypothesis, and contra the Enumerated Alternatives hypothesis, but the difference was not reliable ($\hat{\beta}$=-0.23, z=-1.04). Overall, then, the results for First Mention are most consistent with the Enumerated Alternatives hypothesis, because they show a shift away from next mention of the Goal following Goal prominence.

Turning to whether the continuations evidenced Contrast, the results revealed that despite the presence of a mentioned set of alternative characters, a required choice between Themes, visual scenes that displayed the alternatives and provided labels for ease of mention, truncated context sentences that allowed the participants to assume an alternative-inducing continuation rise in sentence-final position, and highly salient L+H* pitch accents to indicate prosodic contrast, the incidence of Contrast was quite low (Table 2). Only about one-fifth of the continuations expressed some kind of listing of alternative sub-questions, even under these supportive conditions.

Table 2 Mean proportions of Contrast by prominence

	Contrast

Broad prominence	0.16
Goal prominence	0.22
Source prominence	0.19

Our statistical model was set up in identical fashion to the one for First Mention, except that the dependent measure was the presence versus absence of Contrast. It indicated significantly more Contrast with Goal prominence than Broad prominence (Goal prominence: .22, vs. Broad: .16, $\hat{\beta}$=1.91, z=2.09, p<.05), but only a marginal increase in Contrast for Source prominence versus Broad prominence (Source prominence: .19, vs. Broad: .16, $\hat{\beta}$=1.79, z=1.94, p=.053). The relatively weak effect of prominence on Contrast does not appear to be due to perseveration of contrastive responses in the Broad prominence condition from the influence of the two contrastive conditions: the average percentage of continuations with Contrast in the critical conditions was well below 50% ($\hat{\beta}$=-2.99, z=-5.13, p<.01).

Although exact counts are complicated by the syntactic variation described above, Broad prominence conditions tended to express contrast with strongly parallel continuations like (3b) that contrasted two or three of the arguments in the context sentence. Goal prominence continuations followed multiple patterns of showing contrast, but almost always included contrast of the Goal and/or Theme object (*Maria gave Emma a bottle of nice wine and gave Lani a silver platter to celebrate; Michelle presented Grace with a hand-made quilt, while Grace presented Michelle with a framed photo; Tom passed Nick an instruction manual and Andrew went and bought a set of cables*), while Source prominence conditions overwhelmingly included contrast with the Source argument (*Sally handed Cindy a pair of scissors and Christie painted with a set of paints*).

3 Discussion and Conclusion

As in our earlier results, we found significant effects of prosody on co-reference, supporting the general claim that prosodic information[5] influences processing decisions about next mention. Although the strength of the results has varied across studies, the fact that prosodic effects have emerged across different samples of prosodic and syntactic materials and across different laboratories and test populations speaks to its importance as a factor in co-reference.

Notably, the current experiment and our earlier work produced dissimilar patterns of results. Schafer, Takeda, Camp, et al. (2015) found that native English speakers produced more Goal reference with Goal prominence than with Source prominence,[6] suggesting prominence facilitates next mention in a manner consistent with the Salience hypothesis. The current results show a different pattern: the two contrastive conditions were quite similar to each other in proportions of Goal versus Source reference, and Goal reference was significantly *lower* with Goal prominence than with Broad prominence, as predicted by the Enumerated Alternatives hypothesis, and not *higher*, as predicted by the Salience hypothesis. The differential influence of Goal prominence cannot be attributed to variant realization of the critical pitch accents across studies, as in proposals that divide the interpretation of H* versus L+H* pitch accents to discourse-new versus contrastive information (Steedman, 2014), because the current study drew on the same contrastive tokens tested in the previous work (i.e. the same recordings, truncated to remove the original Theme). Instead, the larger discourse situation appears to be critical.

[5] We assume that the referential effects we have found are primarily due to differences in pitch accent patterns across our prominence conditions. However, our conditions also differed in prosodic phrasing; as described above, prominence was realized with L+H* L-H% tunes on the prominent argument. It is not always clear exactly how the prosody varied in the stimuli other researchers have tested, e.g., whether the differences were limited to pitch accents or not.

[6] Broad prominence sentences were not included in this study. It tested 20 critical items in a cross of Goal/Source prominence and grammatical aspect and found similar prominence effects across aspectual conditions.

We speculate that there was little effect of Source prominence on continuations in the current experiment because it was readily interpreted as marking a shift in topic and syntactic subject from the background sentence. Regardless of whether the prosodic prominence was taken to mark the need to access a relatively less accessible discourse entity or the selection of a character from the set of introduced alternatives, it fulfilled an interpretable discourse function that was satisfied by discourse properties within the critical sentence. We assume Goal prominence less readily supported a topic shift from the background to the context sentence, and so more strongly indicated an expression of contrast to the participants, who could then easily realize contrastive alternatives in their continuations given the open-ended task and contextual support for alternatives. In the earlier work, there was less support for contrastive alternatives, as outlined above. In addition, that experiment prompted continuations by supplying a subject pronoun, which we believe favored a search for an antecedent that was salient. Under these circumstances, the general salience-lending properties of the critical accents could hold sway, especially at the point after the pronoun had been encountered.

The relationship between pitch accentuation and information structure in English is complex. On the one hand, focus can project from a pitch-accented element to a larger phrase (Selkirk, 1984, 1995) or be expected in a default location (Büring, 2016). On the other hand, material can bear a pitch accent solely to meet phonological requirements or rhythmic preferences without placing the material in semantic focus (Shattuck-Hufnagel & Turk, 1996), and some categories of pitch accents predominantly convey that the accented material is inferable from the interlocutor's prior knowledge state (Pierrehumbert & Hirschberg, 1990). Our experimental findings further support the view that prosodic contours often support more than one interpretation (Dennison, 2010; Dennison & Schafer, 2017) and must be interpreted with

respect to the larger discourse context (Pierrehumbert & Hirschberg, 1990). It is insufficient to consider simply whether a phrase is pitch-accented or not in the determination of sentence meaning, since the same realization of a pitch accent can lead to different interpretations in different contexts, and different categories and realizations of pitch accents significantly affect meaning. Indeed, in the current experiment, the Source and Goal arguments each received a pitch accent in every prominence condition, yet the different prominence conditions affected continuation choices.

While the current experiment found support for Enumerated Alternatives in the Goal prominence condition, the overall level of Contrast was quite low. It should be emphasized that the low prevalence of Contrast in these results does not mean that participants had no expectation of eventual mention of alternatives in the story. It remains possible that they first offered results, explanations, and so forth of the event described by the context sentence and would have gone on to offer alternative events had they continued more of the story. Nevertheless, the results suggest that as a general rule the participants were not compelled to immediately supply alternatives, as we might expect if sub-questions of a QUD are treated as linguistic relations that should be resolved at the earliest opportunity, or if there is a close grammatical association between the critical tune and the subsequent mention of alternatives. Instead, participants responded to the verb-based bias of transfer-of-possession events and predominantly continued with information about the Goal or end-state. It may be the case that an L+H* L-H% tune is frequently employed when other factors, such as an explicit QUD, establish sub-questions about alternatives, but the tune itself does not create a strong bias for an implicit QUD that supports Enumerated Alternatives.

The current study was exploratory in nature, and much further research will be necessary to support the speculations given here. For example, it would be useful to examine which tunes native speakers most frequently produce, in different discourse situations, to convey the types of continuations and coherence relations considered here. Looking forward, we believe that more attention to the larger discourse context will be critical to disentangling how prosodic form relates to meaning. More generally, we argue that an adequate characterization of prosody's role in discourse processing requires consideration of its representation at multiple levels of linguistic analysis, and of how the processing system interfaces with each of these levels incrementally as production and comprehension proceed. We are eager to see how this area of inquiry, so strongly shaped by Frazier's insights and foundational research, will continue to develop in years to come.

Acknowledgments

We are grateful to all who made it possible for us to join in the celebration of Lyn Frazier and her contributions to our field, including, most prominently, Lyn Frazier herself. This research was supported by a grant to T. Grüter and A. Schafer from the National Science Foundation (BCS-1251450). It was further supported by research assistance from A. L. Blake, Bonnie Fox, Victoria Lee, Wenyi Ling, Ivana Matson, and Maho Takahashi, and helpful comments from the reviewers of this chapter and attendees at Lynschrift18. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Arnold, J. E. (2001). The effect of thematic roles on pronoun use and frequency of reference continuation. *Discourse Processes*, *31*(2), 137-162.

Arnold, J. E. (2010). How speakers refer: the role of accessibility. *Language and Linguistics Compass, 4,* 187-203.

Arnold, J. E., Kaiser, E., Kahn, J. M., & Kim, L. K. (2013). Information structure: linguistic, cognitive, and processing approaches. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(4), 403-413.

Balogh, J. (2003). *Pronouns, prosody, and the discourse anaphora weighting approach* (Unpublished doctoral dissertation). University of California, San Diego.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48. doi:10.18637/jss.v067.i01.

Beckman, M., & Ayers Elam, G. (1997). Guidelines for ToBI labeling, version 3. Unpublished manuscript. Department of Linguistics, Ohio State University, Columbus.

Braun, B. & Tagliapietra, L. (2011). On-line interpretation of intonational meaning in L2. *Language and Cognitive Processes, 26,* 224-235.

Breen, M., Fedorenko, E., Wagner, M., & Gibson, E. (2010). Acoustic correlates of information structure. *Language and Cognitive Processes*, *25*(7), 1044–1098.

Büring, D. (2016). *Intonation and meaning*. Oxford, England: Oxford University Press.

Calhoun, S. (2010). The centrality of metrical structure in signaling information structure: a probabilistic perspective. *Language*, *86*(1)*,* 1-42.

Carlson, K. (2002). *Parallelism and prosody in the processing of ellipsis sentences.* New York: Routledge.

Carlson, K. (2009). How prosody influences sentence comprehension. *Language and Linguistics Compass, 3/5,* 1188-1200.

Carlson, K., & Harris, J. A. (2018). Zero-Adjective contrast in much-less ellipsis: the advantage for parallel syntax. *Language, Cognition and Neuroscience*, *33*(1), 77-97.

Carlson, K., Dickey, M., Frazier, L., Clifton, C. (2009). Information structure expectations in sentence comprehension. *The Quarterly Journal of Experimental Psychology,* 62, 114-139.

Colonna, S., Schimke, S. & Hemforth, B. (2015). Different effects of focus in intra- and inter-sentential pronoun resolution in German. *Language, Cognition and Neuroscience, 30,* 1306-1325.

Constant, N. (2012). English rise-fall-rise: a study in the semantics and pragmatics of intonation. *Linguistics and Philosophy*, *35*, 407-442.

Cowles, H., Walenski, M., & Kluender, R. (2007). Linguistic and cognitive prominence in anaphor resolution: topic, contrastive focus and pronouns. *Topoi,* 26, 3-18.

Dahan, D. (2015). Prosody and language comprehension. *Wiley Interdisciplinary Reviews: Cognitive Science*, *6*(5), 441–452.

Dahan, D., Tanenhaus, M., & Chambers, C. G. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language, 47,* 292-314.

Dennison, H. (2010). Processing implied meaning through contrastive prosody (Unpublished doctoral dissertation). University of Hawaii, Manoa.

Dennison, H. & Schafer, A. J. (2017). Processing intonationally implicated contrast versus negation in American English. *Language and Speech, 60,* 174-199.

Fraundorf, S. H., Watson, D. G., & Benjamin, A. S. (2010). Recognition memory reveals just how CONTRASTIVE contrastive accenting really is. *Journal of Memory and Language, 63,* 367-386.

Frazier, L. (1987). Syntactic processing: evidence from Dutch. *Natural Language & Linguistic Theory*, *5*(4), 519-559.

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, *17*(3), 311-347.

Grüter, T., Rohde, H., & Schafer, A. J. (2017). Coreference and discourse coherence in L2: The roles of grammatical aspect and referential form. *Linguistic Approaches to Bilingualism, 7, 199-229.*

Husband, E. M., & Ferreira, F. (2016). The role of selection in the comprehension of focus alternatives. *Language, Cognition and Neuroscience, 31,* 217-235.

Ito, K., & Speer, S. R. (2008). Anticipatory effects of intonation: eye movements during instructed visual search. *Journal of Memory and Language, 58,* 541–73.

Itzhak, I., & Baum, S. R. (2015). Misleading bias-driven expectations in referential processing and the facilitative role of contrastive accent. *Journal of Psycholinguistic Research, 44,* 623–650.

Jackendoff, R. S. (1972). *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

Kaiser, E. (2010). Investigating the consequences of focus on the production and comprehension of referring expressions. *International Review of Pragmatics*, *2*(2), 266-297.

Kaiser, E. (2011). Focusing on pronouns: Consequences of subjecthood, pronominalisation, and contrastive focus. *Language and Cognitive Processes, 26,* 1625-1666.

Kehler, A. (2002). *Coherence, reference, and the theory of grammar.* Stanford, CA: CSLI Publications.

Kehler, A., Kertz, L., Rohde, H., & Elman, J. L. (2008). Coherence and coreference revisited. *Journal of Semantics*, *25,* 1-44.

Kim, C. S., Gunlogson, C., Tanenhaus, M. K., & Runner, J. T. (2015). Context-driven expectations about focus alternatives. *Cognition*, *139*, 28-49.

Kurumada, C., Brown, M., Bibyk, S., Pontillo, D. F., & Tanenhaus, M. K. (2014). Is it or isn't it: Listeners make rapid use of prosody to infer speaker meanings. *Cognition, 133*, 335–342.

Pierrehumbert, J., & Hirschberg, J. (1990). The meaning of intonational contours in interpretation of discourse. In P. Cohen, J. Morgan, & M. Pollack (Eds.), *Intentions in communication* (pp. 271-311). Cambridge, MA: MIT Press.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Roberts, C. (2012). Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics & Pragmatics, 5,* 1-69.

Rohde, H. & Kehler, A. (2014). Grammatical and information-structural influences on pronoun production. *Language, Cognition, and Neuroscience, 29,* 912-927.

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, *1*, 75-116.

Schafer, A. J., Carter, J., Clifton, C., & Frazier, L. (1996). Focus in relative clause construal. *Language and Cognitive Processes*, 11, 135-163.

Schafer, A. J., Takeda, A., Camp, A., Rohde, H., & Grüter, T. (2015). Effects of contrastive intonation and grammatical aspect on processing coreference in Mainstream American English. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th*

International Congress of the Phonetic Sciences, Glasgow, Scotland: The University of Glasgow. ISBN 978-0-85261-941-4. Paper number 463. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0463.pdf

Schafer, A. J., Takeda, A., Rohde, H., & Grüter, T. (2015, November). Mapping prosody to reference in L2. Poster presented at the 40th Boston University Conference on Language Development, Boston, MA.

Schumacher, P. B., Backhaus, J., & Dangl, M. (2015). Backward- and forward-looking potential of anaphors. *Frontiers in Psychology, 6,* 1746.

Selkirk, E. O. (1984). *Phonology and syntax*. Cambridge, MA: MIT Press.

Selkirk, E. O. (1995). Sentence prosody: Intonation, stress and phrasing. In J. Goldsmith (Ed.), *Handbook of phonological theory* (pp. 550–569). Oxford, England: Basil Blackwell.

Shattuck-Hufnagel, S., & Turk, A.E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, *25*(2), 193-247.

Steedman, M. (2014). The surface-compositional semantics of English intonation. *Language*, *90*(1), 2-57.

Stevenson, R. J., Crawley, R. A., & Kleinman, D. (1994). Thematic roles, focusing and the representation of events. *Language and Cognitive Processes, 9*, 519-548.

Tomioka, S. (2010). Contrastive topics operate on speech acts. In C. Féry & M. Zimmerman (Eds.), *Information structure: Theoretical, typological, and experimental perspectives*. (pp. 115-138). Oxford, England: Oxford University Press.