Effects of contrastive intonation and grammatical aspect on processing coreference in Mainstream American English

Amy J. Schafer[1], Aya Takeda[1], Amber Camp[1], Hannah Rohde[2], and Theres Grüter[1]

[1]University of Hawaii and [2]University of Edinburgh
aschafer@hawaii.edu

ABSTRACT

Coreference choices are influenced by multiple factors, including information structural categories such as topic and focus. These information structural categories can be indicated by intonation, yet few studies have investigated how intonation affects subsequent choices for coreference. Using a story continuation experiment with aurally presented stimuli, we show that the location of contrastive focus in Mainstream American English significantly affects the preferred referent for the subject of the next sentence in a short discourse.

Keywords: discourse, focus, intonation, processing, reference

1. INTRODUCTION

Intonation provides critical cues to the information structure of a sentence [e.g., 16]. Morphosyntactic markings of information structure (e.g., clefting) have robust effects on coreference [2], making it easy to postulate a causal chain in which intonational cues affect information structure, which in turn guides coreference. Yet the exploration of how intonation affects coreference has been limited, both at the empirical level (*Which specific intonational contours facilitate which coreference choices?*) and in terms of hypotheses about the relevant processing mechanisms (*How does intonation change a referent's status in the comprehender's memory or in his/her expectations about upcoming discourse?*). Moreover, previous research suggests opposing predictions for how intonation could affect subsequent coreference. We therefore examine how contrastive focus placement in Mainstream American English sentences affects coreference patterns for the sentential subject of the next sentence in a short discourse.

2. BACKGROUND

An important component of producing and comprehending discourse is the management of reference to relevant entities. Reference back to previously introduced information is one of the primary strategies speakers use to establish discourse cohesion and coherence. For example, language producers tend to maintain the topic of the discourse across sentences [7], and typically mark reference to the established topic with reduced expressions, such as the choice of a pronoun instead of a name or full noun phrase [15].

Psycholinguistic research in this domain has identified multiple factors that work together rapidly and incrementally to guide reference resolution [2]. One factor, which we exploit here as a control factor, is event structure, as marked by grammatical aspect. A series of experiments using written English have tested sentence contrasts like that in (1) [14, 20, 24]. Sentences with perfective aspect (*gave*) describe an event as completed, whereas ones with imperfective aspect (*was giving*) describe it as ongoing. This contrast between completed and ongoing events leads to different biases for how the discourse will most naturally unfold, i.e. expectations about upcoming coherence relations [17, 19].

> (1) Patrick [gave/was giving] Ron a warm towel.
> (2) He said "thanks, Patrick!"
> (3) He wanted Ron to dry off.

Completed events have a salient end state and are hence compatible with discussion of what happened next. Therefore, in a sentence like (1) that describes a transfer of possession of the towel from Patrick to Ron, a likely continuation might be the sentence in (2), which describes a plausible result of the giving event. Furthermore, a description of what happened next is likely to start with mention of the individual associated with the end state of the transfer event, here the Goal, Ron.

Ongoing events, on the other hand, are more likely to encourage further elaborations on the details of the event and explanations for it. These types of coherence relations are likely to begin with mention of the individual associated with the start state of a transfer event, the Source. This coherence-driven bias to elaborate or explain the transfer event thus leads to more sentences like (3) that begin with reference to Patrick, the Source of the giving event. Previous research has established that participants who are given written stimuli like (1) and asked to

write the next sentence in an imagined story will shift in the proportion of initial Source versus initial Goal reference depending on the grammatical aspect of the stimulus sentence. Perfective aspect favors continuations that begin with Goal reference, wherein the previous Goal argument becomes the sentential subject/topic of the next sentence, and imperfective aspect shifts the bias the other way, toward initial Source-reference continuations.

Turning to intonation, multiple studies have established that comprehenders can rapidly use contrastive intonation to establish discourse contrast, [e.g., 11], and that intonationally focused material is more easily retrieved from memory for subsequent recognition or recall [e.g., 5, 12]. Studies on ellipsis constructions (e.g., *The captain talked to the co-pilot, but we couldn't find out who else*), have found significant effects of morphosyntactic cues to information structure, as well as significant effects of contrastive intonation, with both types of cues favoring a focused antecedent [8, 9, 13].

Studies on the effects of contrastive intonation on cross-sentence coreference have been surprisingly rare. One series of experiments tested coreferential preferences for object pronouns and found mixed results; accentuation increased coreference to the accented material in some of the experiments but not others [3]. Unfortunately, characterization of the intonation used for these experiments was limited.

Another study manipulated the intonational realization of the focused material in *it*-clefts, comparing more contrastive productions (realized with L+H* accents) to more neutral ones (mostly L+H* accents with smaller pitch excursions than in the contrastive condition) and found no significant effect of intonation, although the more contrastive intonation may have enhanced the general effect of clefting on reference [10]. A related experiment in the study found that discourse topics, sentence topics, and material contrastively focused by *it*-clefts all had similar, facilitative, effects on antecedent availability (see also [1, 18]). Taken together, these studies are suggestive that intonationally-marked contrastive focus could lead to increased reference to the focused material, but they are far from conclusive.

The most closely related research to the experiment we present below may come from work on Korean. A set of experiments using Korean versions of sentences like (1), and a similar story continuation task as the earlier aspect studies but with spoken stimuli, found that both aspect and intonationally-marked contrast had significant effects on coreference [21, 22]. Strikingly, the specific effect of contrastive intonation depended on whether it was realized on the Source argument or

the Goal. Contrastive intonation on the Source argument increased coreference to it compared to sentences with Broad focus. However, contrastive intonation on the Goal argument resulted in *decreased* coreference to the Goal compared to Broad focus stimuli, so that there was a higher proportion of Source reference with Goal focus than with Source focus. Instead of switching the discourse topic from the Source to a contrastively focused Goal, the Korean participants increased the proportion of sentential subjects with reference to the Source, often mentioning an alternative Goal argument (e.g., the equivalent of *Patrick gave RON a towel. He (=Patrick) didn't give AMANDA one.*). These findings align with ERP results showing that Japanese speakers prefer to maintain an established topic when the discourse establishes contrast [25].

It is possible that English could show a similar pattern to Korean. Nevertheless, Korean and English differ in a number of critical respects, including their intonational patterns, their linguistic mechanisms for marking focus, and their pronominal inventories, so it is also reasonable to predict that the Korean findings would not extend to English.

3. EXPERIMENT

Our experiment aimed to replicate previous effects of aspect in a test with spoken English stimuli, and evaluate how intonationally marked contrastive focus affects coreference. Would contrastive focus on an argument increase coreference to it, so that coreference (uniformly) echoes the contrastive focus location? Or would the effect of contrastive focus on coreference depend on whether the focused argument was already the discourse topic, as in the Korean findings reviewed above, so that Goal focus would *decrease* Goal coreference and *increase* Source coreference?

3.1. Participants

Forty-eight native speakers of English from the University of Hawaii community participated; an additional 4 were tested but eliminated from analysis for excessive errors (3) or language disability (1). Participants were rewarded for their time with either a small amount of course credit or a gift card worth $10.

3.2. Materials

The critical stimuli consisted of 20 sentences describing transfer-of-possession events, as in (1), given above. In each critical sentence, two people of the same gender were mentioned. One was the Source of the transfer event, and always appeared as
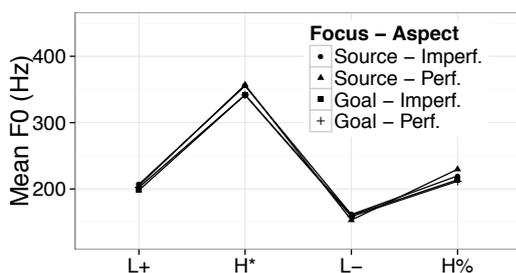
the sentential subject. The other was the Goal, and always appeared as the indirect object in a double-object construction, followed by the Theme object. Two versions of each sentence were constructed, with perfective versus imperfective aspect. An additional 40 filler sentences contained a range of other constructions.

The stimuli were recorded by a native speaker of English trained in the ToBI intonational system for Mainstream American English [4]. Each critical sentence was recorded with contrastive focus on the Source, contrastive focus on the Goal, and, for comparison purposes and as stimuli for a related experiment, Broad focus. All stimuli were produced in a style appropriate for non-native speakers of English, in preparation for a related experiment. The stimuli employed a "clear" style of speech often employed in psycholinguistic stimuli and speech directed at classes, radio audiences, non-native speakers and the like, in which each content word in the stimuli was accented. Therefore, the focus manipulation was realized through the type of pitch accent and the prosodic phrasing of the materials, not via a simple contrast between the presence or absence of pitch accents.

3.2.1. Intonational analysis

The intonation of each critical token and their Broad focus counterparts was transcribed in MAE-ToBI by a trained coder, and then verified by a second trained coder. The few cases of disagreement between the coders were resolved by discussion. Except for a few minor variations, all stimuli within a focus treatment followed the same pattern. Contrastive focus was marked with a contrastive accent and a continuation rise (L+H* L-H%) [23] on the contrastively focused region (Figure 1). All remaining content words were marked with L*+H accents on Source arguments, H* accents on Goal arguments and Themes, and H*, H+!H*, or L*+H accents on Verb regions, with downstepped realizations for all phrase-medial accents in all regions.

Figure 1: Mean F0 values for each tone in the L+H* L-H% contours, for the contrastively focused region in each critical condition.

3.2.2. Acoustical analysis

The Source, Verb, Goal, and Theme regions of each token were analysed for duration and minimum and maximum F0 using an F0 range of 125 to 450 Hz and the default pitch settings of Praat [6]. Critical F0 locations were hand-corrected for any tracking errors. Contrastively focused Source and Goal regions were marked for the realization point of each tone in the L+H* L-H% contour, and all other pitch accent locations were marked at the most representative point in the accented vowel.

Repeated measures ANOVAs verified that contrastively focused regions were realized with greater duration (speech + following silence, Figure 2) and F0 excursion (F0max-F0min, Figure 3) than their counterparts in other conditions. Mean F0 values for the L+H* L-H% contrastive contours (in the Source region for the Source focus condition and the Goal region for Goal focus) were nearly identical across critical conditions (Figure 1). Perfective sentences had shorter durations and smaller F0 excursions in the Verb region than their Imperfective counterparts, as expected given the additional material (*was ...-ing*) for the Imperfectives.

Figure 2: Mean duration (ms) for each region (across panels) for Source, Broad, and Goal focus by aspect, with 95% confidence intervals.
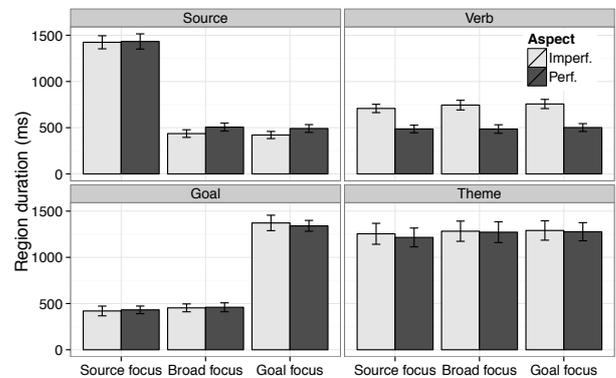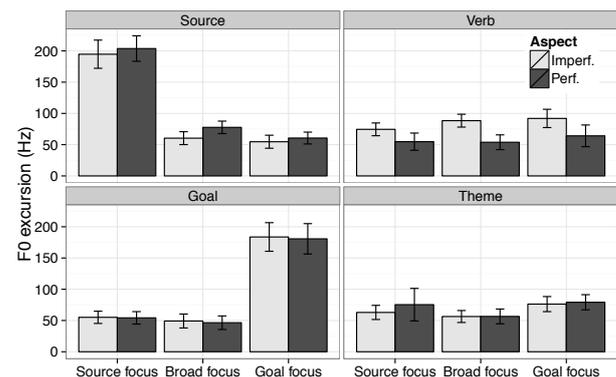

Figure 3: Mean F0 excursion (Hz) for each region (across panels) for Source, Broad, and Goal focus by aspect, with 95% confidence intervals.

3.3. Procedure and design

The experiment employed a story continuation task, in which participants completed a series of short discourses. Each trial began with the presentation of a recorded sentence, played over headphones at a comfortable volume. After its conclusion, a computer monitor displayed a pronoun prompt, along with a text box. Participants were asked to type the first natural continuation that occurred to them into the text box, starting with the prompted pronoun. They then moved on to the next trial.

The experiment started with 6 practice items, followed by a pseudorandomized mix of the 20 critical items and 40 fillers. The 4 experimental conditions were comprised of a 2 x 2 design crossing aspect (imperfective, perfective) with contrastive focus location (Source focus, Goal focus). These 4 conditions were counterbalanced across 4 lists, so that each participant received 5 critical items in each condition, and all items were tested in all conditions across the set of participants. Following the main experimental block, an additional block presented 12 critical items (3 per condition) in a forced-choice focus-location selection task, used to verify that all participants perceived the focus manipulation. Mean accuracy on this focus-check was 98.6%. The entire experiment took less than one hour to complete.
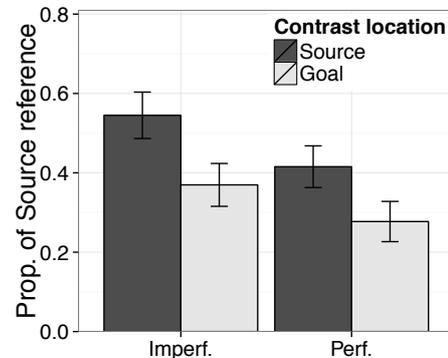
3.4. Results

Each typed response was coded by two trained annotators, who separately evaluated whether the pronoun that began the continuation referred to the Source or Goal of the previous sentence. Other types of completions, such as incomplete responses, reference outside of the previous sentence, and ambiguous reference were eliminated from analysis, along with any continuation on which the two coders disagreed. This procedure removed 10.5% of the total data. The remaining data were then fit to a mixed-effects logistic regression model that incorporated aspect and contrastive focus location as fixed effects and participants and items as random effects. For each random effect the model included intercepts and slopes for the two fixed effects.

Replicating previous results with written materials [14, 20, 24] and as shown in Figure 4, there was a significant effect of aspect, with the expected higher proportion of Source reference for imperfective sentences than for perfective ones ($B = 0.67$, se = 0.23, z = 2.89, p < .005).

The results also revealed a significant effect of contrastive focus location ($B = -.95$, se = 0.28, z = -3.37, p < .001), which did not interact at a significant level with aspect (p = .92). Coreference echoed the contrastive focus location, so that Source reference was greater in the Source focus conditions (dark bars) than in the Goal focus ones (light bars).

Figure 4: Proportion of Source reference by contrastive focus location and aspect, with 95% confidence intervals corrected for repeated measures.

4. GENERAL DISCUSSION

Both aspect and contrastive intonation had significant effects on coreferential patterns, extending previous effects of aspect to spoken English stimuli, and replicating the general finding that multiple factors influence coreference.

Unlike the findings for Korean [22], contrastive intonation increased reference to the focused argument regardless of its position or topic status. Further research will be necessary to investigate if differences between the two languages in intonation or other mechanisms for encoding discourse and information structure (e.g., use of argument drop, scrambling) explain the dissimilar patterns.

The current results show highly significant effects of intonation on coreference. It is likely that the present effects are stronger than previous ones because the intonational contrasts tested here were more salient than in earlier stimuli. Additional studies with more subtle intonational manipulations would be valuable for defining which types of contrasts are necessary or sufficient for shifting coreference. Such studies could also sharpen our understanding of the exact mechanism connecting intonation to coreference. For example, to what extent do comprehenders draw on distinctions encoded at phonetic, phonological, or information-structural levels when responding to contrastive focus? The present results are consistent with multiple mechanisms, including information-structural distinctions associated with predictions for upcoming material [14], and saliency distinctions associated with stronger memory traces for reactivation of earlier material [12]. Current investigations in our laboratory are exploring these and related questions.

5. REFERENCES

[1] Arnold, J. 1998. *Reference form and discourse patterns.* Unpublished doctoral dissertation. Stanford University.

[2] Arnold, J. 2010. How speakers refer: The role of accessibility. *Lang. and Linguistic Compass* 4, 187-203.

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

[4] Beckman, M., Ayers, G. E. 1997. Guidelines for ToBI labeling, version 3. Unpublished manuscript. Ohio State University.

[5] Birch, S., Garnsey, S. 1995. The effect of focus on memory for words in sentences. *J. of Mem. and Lang.* 34, 232-267.

[6] Boersma, P., Weenink, D. 2014. Praat: Doing phonetics by computer (Version 5.3.65) [Computer program]. www.praat.org

[7] Brennan, S. 1995. Centering attention in discourse. *Lang. and Cog. Proc.* 10, 137-167.

[8] Carlson, K. 2001. The effects of parallelism and prosody on the processing of gapping structures. *Lang. and Speech* 44, 1-26.

[9] Carlson, K., Dickey, M., Frazier, L., Clifton, C. 2009. Information structure expectations in sentence comprehension. *The Quarterly J. of Exp. Psych.* 62, 114-139.

[10] Cowles, H., Walenski, M., Kluender, R. 2007. Linguistic and cognitive prominence in anaphor resolution: Topic, contrastive focus and pronouns. *Topoi* 26, 3-18.

[11] Dahan, D., Tanenhaus, M., Chambers, C. 2002. Accent and reference resolution in spoken-language comprehension. *J. of Mem. and Lang.* 47, 292-314.

[12] Fraundorf, S., Watson, D., Benjamin, A. 2010. Recognition memory reveals just how *contrastive* contrastive accenting really is. *J. of Mem. and Lang.* 63, 367-386.

[13] Frazier, L., Clifton, Jr., C., Carlson, K. (2007). Focus and VP ellipsis. *Lang. and Speech,* 50, 1-21.

[14] Grüter, T., Rohde, H., Schafer, A. 2014. The role of discourse-level expectations in non-native speakers' referential choices. In: Orman, W., Valleau, M. J. (eds), *Proc. 38th BUCLD.* Somerville, MA: Cascadilla, 179-191.

[15] Gundel, J., Hedberg, N., Zacharski, R. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69, 274–307.

[16] Gussenhoven, C. 2004. *The phonology of tone and intonation.* Cambridge: Cambridge University Press.

[17] Hobbs, J. R. 1979. Coherence and coreference. *Cog. Sci.* 3, 67–90.

[18] Kaiser, E. 2011. Focusing on pronouns: Consequences of subjecthood, pronominalisation, and contrastive focus. *Lang. and Cog. Proc.* 26, 1625-1666.

[19] Kehler, A. 2002. *Coherence, reference, and the theory of grammar.* Stanford, CA: CSLI Publications.

[20] Kehler, A., Kertz, L., Rohde, H., Elman, J. 2008. Coherence and coreference revisited. *J. of Semantics* 25, 1-44.

[21] Kim, K., Grüter, T., Schafer, A. J. 2013. Effects of event-structure and topic/focus-marking on pronoun reference in Korean. 26th Annual CUNY Conference on Human Sentence Processing, Columbia, SC.

[22] Kim, K., Grüter, T., Schafer, A. J. 2014. Effects of morphological and prosodic focus cues on topic maintenance in Korean. 27th Annual CUNY Conference on Human Sentence Processing, Columbus, OH.

[23] Pierrehumbert, J., Hirschberg, J. 1990. The meaning of intonational contours in the interpretation of discourse. In: Cohen, P. R., Morgan, J., Pollack, M. E. (eds), *Intentions in Communication.* Cambridge, MA: MIT Press, 271-323.

[24] Rohde, H., Kehler, A., Elman, J. 2006. Event structure and discourse coherence biases in pronoun interpretation. In R. Sun (ed), Proc. 28th Cog. Sci. Lawrence Erlbaum Associates. Mahwah, NJ, 617–22.

[25] Wang, L., Schumacher, P. 2013. New is not always costly: Evidence from online processing of topic and contrast in Japanese. *Frontiers in Psychology*, 4.