



HHS Public Access

Author manuscript

J Exp Psychol Learn Mem Cogn. Author manuscript; available in PMC 2022 September 01.

Published in final edited form as:

J Exp Psychol Learn Mem Cogn. 2021 September ; 47(9): 1494–1526. doi:10.1037/xlm0000986.

Eye see what you're saying: Contrastive use of beat gesture and pitch accent affects online interpretation of spoken discourse

Laura M. Morett,

Dept. of Educational Studies in Psychology, Research Methodology, and Counseling University of Alabama

Scott H. Fraundorf,

Dept. of Psychology, Learning Research and Development Center University of Pittsburgh

James C. McPartland

Dept. of Child Psychiatry Yale University

Abstract

Cues to prominence such as beat gesture and contrastive pitch accent play an important role in constraining what is remembered. However, it is currently unclear how beat gesture affects online discourse processing alone and in combination with contrastive accenting. Using an adaptation of the visual world eye-tracking paradigm, we orthogonally manipulated the presence of these cues and their felicity (match) with contrast within local (sentence-level) and global (experiment-level) referential contexts. In Experiment 1, in which beat gesture and contrastive accenting were always globally felicitous with the context of filler referring expressions, beat gesture increased anticipation of both target and competitor referents of locally infelicitous critical referring expressions differing in color and shape, whereas contrastive accenting hindered resolution of these expressions. In Experiment 2, in which beat gesture and contrastive accenting were always globally infelicitous with the context of filler referring expressions, beat gesture increased anticipation of both target and competitor referents of locally felicitous critical referring expressions contrasting in color, whereas contrastive accenting did not affect their interpretation. Taken together, these findings indicate that local and global felicity of cues to prominence with contrast affects their interpretation during online spoken discourse processing.

Keywords

Beat gesture; pitch accent; linguistic contrast; visual world; eyetracking

To comprehend discourse successfully, it is necessary to establish how entities within a discursive context are related to one another. One way in which such entities may be related is via contrast, which refers to a contradiction between relevant features (e.g., *The meeting isn't on Friday; it is on Monday*; Myhill & Xing, 1996). Although contrast can be discerned semantically after the fact based on salient contradictions between entities

(Gotzner & Spalek, 2019), it can also be anticipated based on potentially diagnostic features within the communicative environment, which we term *cues* (e.g., font, speech, gestural emphasis). Cues facilitate discourse comprehension and strengthen mental representations of propositional relations between contrasting entities insofar as they convey contrast reliably (Fraundorf et al., 2010, 2012, 2013; Lee & Snedeker, 2016; Sanford et al., 2006). Thus, cues that consistently co-occur with contrastive information enhance its processing to a greater degree than cues that inconsistently occur with contrastive information (Grodner & Sedivy, 2011; Roettger & Franke, 2019; Roettger & Rimland, 2020; Ryskin et al., 2019). Crucially, the effects of cue reliability are interactive, such that highly reliable cues to contrast influence the effects of less reliable cues to contrast on memory for contrastive information to a much greater degree than less reliable cues influence the effects of highly reliable cues—at least in offline memory for contrastive information (Morett & Fraundorf, 2019). At present, however, it is unclear whether reliability of different contrast cues affects online processing of contrastive information in a similarly interactive manner, and, if so, whether its effects are constrained to contrast resolution or whether they also influence prediction of upcoming contrasts.

In the current research, we examine how—and when—online interpretation of contrastive information is affected by the reliability with which pitch accenting and beat gesture indicate contrast both within individual referring expressions (local felicity) as well as the broader communicative context (global felicity). To do so, we leverage an adaptation of the visual world paradigm including video (Silverman et al., 2010) so that we can examine interactive effects of pitch accenting and beat gesture on online spoken discourse processing using carefully controlled experimental stimuli. In doing so, we provide insight into the individual and combined contributions of multimodal cues to prominence in the prediction and resolution of contrast in particular, as well as online processing and representation of inter-entity relations more generally. Moreover, we compare these effects across different discursive contexts in which these cues differ in felicity.

Pitch Accent Interpretation

To convey the discourse status of referents, talkers often use *pitch accents*, phonological constructs realized acoustically with changes in the fundamental frequency (F_0), duration or intensity that increase the prominence of referring expressions (for a review, see Ladd, 1996). In one influential transcription scheme for pitch accents, ToBI (Tones and Breaks Index), two of the most common pitch accents in English are referred to as H* and L+H*. The H* pitch accent consists of a high pitch target with F_0 high in the talker's range, whereas the L+H* pitch accent consists of an initial low pitch (L) followed by a sharp rise to a high target on the accented syllable (H*). The H* pitch accent is used broadly with any kind of new information, whereas the L+H* pitch accent is often used specifically with information that contrasts directly with other parts of a discourse (Pierrehumbert & Hirschberg, 1990). An example of a discourse with both of these pitch accents follows:

- (1a) [S1] Which of my friends is coming to the party on Saturday?
 (1b) [S2] I think Jenny is coming.

(2a) [S1] Did you say Ashley is coming?

(2b) [S2] No, I said JENNY is coming.

In (1b), *Jenny* is new information and would likely have a presentational (H*) accent. In (2b), however, *Jenny* contrasts with a referent previously mentioned in (2a), *Ashley*, and would likely have a contrastive (L+H*) accent.

Broad evidence suggests that, in language comprehension, contrastive pitch accenting directs listeners' attention to a contrast between the current referent and a previously-mentioned referent, whereas presentational pitch accenting directs listeners' attention to new referents more generally (for a review, see Gotzner & Spalek, 2019). For instance, offline memory for referents with contrastive accenting is superior to memory for referents with presentational accenting, particularly when a salient contrasting referent must be rejected (Fraundorf et al., 2010, 2012; Lee & Snedeker, 2016; Sanford et al., 2006). Moreover, in online language comprehension, contrastive pitch accenting facilitates rejection of objects contrasting with a spoken referent (e.g., [satellite] *dish*, which is an alternative to *antenna*; Braun & Tagliapietra, 2010; Husband & Ferreira, 2016). A similar effect is evident in eye-tracking studies, in which contrastive accenting promotes attention to objects that contrast with a previously-mentioned referent (Ito et al., 2012; Ito & Speer, 2008; Kurumada et al., 2014; Watson et al., 2008; Weber et al., 2006). An example particularly relevant to the design of the present investigation is Experiment 2 of Ito and Speer (2008), which examined how looking patterns differed when pitch accent was manipulated on the adjective of continuation sentences in pairs such as those in 3a-4b below:

(3a) Hang the green drum...Now, hang the blue_{L+H*} drum.

(3b) Hang the green drum...Now, hang the blue H* drum.

(4a) Hang the green drum...Now, hang the blue H* angel.

(4b) Hang the green drum...Now, hang the blue_{L+H*} angel.

In this experiment, pitch accenting was either locally felicitous with contrast (or lack thereof) in adjectives (3a, 4a), neutral (3b), or locally infelicitous the absence of contrast purely in adjectives (4b). When the target referent contrasted in color with the context referent (e.g., *green drum* → *blue drum*), contrastive accenting on the color word of the continuation referring expression facilitated processing relative to presentational accenting, leading listeners to fixate the target more quickly. Moreover, when the target referent differed from the context referent in both color and shape (e.g., *green drum* → *blue angel*), infelicitous contrastive pitch accenting on the color word of the continuation referring expression hindered processing relative to felicitous presentational pitch accenting, leading listeners to initially incorrectly fixate an object contrasting in color with the context referent (*blue drum*). This finding demonstrates that listeners interpret contrastive pitch accenting as indicating referential contrast during online language processing.

Beat Gesture Interpretation

Another cue often used to convey prominence in discourse is beat gesture. According to the most widely-used taxonomy of gesture (McNeill, 1992, 2005), *beat gesture* consists of

simple rhythmic body movements produced concurrently with speech. Prototypically, beat gesture takes the form of punctate downward hand flicks, but it can be articulated using other parts of the body (e.g., finger movements, head nods, foot taps), in other orientations (e.g., horizontal, oblique, curved), and with multiple components (Shattuck-Hufnagel & Ren, 2018). Beat gesture, like pitch accenting, is often used to emphasize important words or phrases in speech, serving as a “gestural yellow highlighter” (McNeill, 2006).

Findings concerning beat gesture’s impact on memory for speech are mixed. On the phonological level, viewing beat gestures in tandem with vowel nuclei facilitates subsequent discrimination between minimal pairs of second-language (L2) words differing in vowel length (Hirata et al., 2014). On the lexical level, producing beat gesture during L2 word learning also predicts L2 word repetition (Morett, 2014). And, beat gesture presented concurrently with auditory words in a decontextualized list enhances memory for those words in adults in both L1 and L2 (Levantinou & Navarretta, 2016), but not for children (So et al., 2012). On the sentence level, viewing beat gesture fails to enhance memory for entire sentences (Biau et al., 2015; Biau & Soto-Faraco, 2013) and may even decrease memory for sentences in some cases in adults (Feyereisen, 2006), although there is some evidence that it may enhance memory for sentences in young children (Vilà-Giménez et al., 2019). On the discourse level, viewing beat gesture fails to enhance narrative comprehension in children (Macoun & Sweller, 2016), but it enhances children’s—though not adults’—memory for spatial route directions (Austin & Sweller, 2014). The range of designs and levels upon which beneficial effects were observed suggests that viewing beat gesture may enhance phonological contrast perception and lexical memory but that it may not enhance sentence and discourse comprehension and memory, and that its effect on language processing is not as definitive as that of pitch accent.

Beat Gesture and Contrastive Accenting as Cues to Contrast

In natural conversation, cues such as pitch accent and beat gesture are often produced in combination both with each other and with key information, and comprehenders must deduce how to choose, weight, and combine these cues. For example, a talker may emphasize a point using both contrastive pitch accenting and beat gesture, and comprehenders must infer the importance of this point from the co-occurrence of these cues.

How comprehenders weight one of these cues may affect how they weight the other because beat gesture and pitch accent are closely related in both perception and production. During speech production, the points of maximum extension of beat gesture (apices) co-occur with the F_0 peaks of pitch accented words (Leonard & Cummins, 2011), demonstrating that beat gesture and pitch accent are closely related in timing. In perception, viewing beat gesture in conjunction with stressed syllables facilitates non-native talkers’ production of difficult-to-pronounce words by reinforcing metric patterns visually (Gluhareva & Prieto, 2017).

Some evidence indicates that words accompanied by beat gesture are produced with higher vowel formants (F_2 and F_3) and are more likely to be perceived as pitch accented than

words unaccompanied by beat gesture (Krahmer & Swerts, 2007). However, other evidence indicates that beat gesture production has no effect on either articulatory (vocalic target) or acoustic (intensity, duration, F_0) correlates of contrastive pitch accenting for native talkers (Roustan & Dohen, 2010a). One possible explanation for this discrepancy is that beat gesture was always accompanied by pitch accent in Roustan and Dohen (2010a), whereas beat gesture was sometimes unaccompanied by pitch accent in Krahmer and Swerts (2007). Considered together, these results suggest that, in light of beat gesture's and pitch accent's similar functions, the absence of pitch accent in a context in which it sometimes—or always—co-occurs with beat gesture may influence how beat gesture is interpreted. Thus, these results highlight the importance of cue reliability with respect to other cues serving similar functions on cue processing, a point we revisit below.

Nevertheless, this work broadly suggests that beat gesture is interpreted as a cue to prominence when it consistently occurs in conjunction with focused information more generally. However, it is less clear whether it is interpreted as a cue specifically to *contrast* when it consistently occurs in conjunction with contrastive information. In much of this prior work, no specific contrasts were present, and beat gesture was examined in relation to presentational pitch accenting, which indicates prominence, but not in relation to contrastive pitch accenting, which indicates contrast (Igalada et al., 2017; Krahmer & Swerts, 2007; Kushch et al., 2018; Leonard & Cummins, 2011; Roustan & Dohen, 2010a, 2010b). Some work has examined how deictic (pointing) gesture relates to stress on contrasting syllables (Esteve-Gibert & Prieto, 2013; Rusiewicz et al., 2013, 2014) and lexical items (Krahmer & Swerts, 2007; Roustan & Dohen, 2010a, 2010b), but this work focused on the temporal relationship between deictic gesture and stress when both cues are produced on demand rather than *whether* these cues co-occur and *how* they are interpreted; further, this work concerns deictic gesture, not beat gesture. In the current work, we more directly test whether beat gesture encourages contrastive interpretation of referring expressions by virtue of its local felicity with specific contrasts, as opposed to highlighting more general differences between key modifiers and nouns. (See Yap, So, Yap, Tan, & Teoh, 2011 for related evidence that iconic gesture semantically primes speech interpretation.)

One possibility is that beat gesture may need to be accompanied by contrastive accenting to convey contrast. To date, most work examining beat gesture as a cue to contrast in relation to contrastive information and contrastive accenting indicates that, when beat gesture and contrastive accenting occur conjointly with selected alternatives from a contrastive pair, subsequent memory for these selected alternatives is enhanced, suggesting that these cues conjointly enrich representations of contrastive information (Kushch & Prieto, 2016; Llanes-Coromina et al., 2018). However, beat gesture always varied concurrently with pitch accenting in these studies (Kushch & Prieto, 2016; Llanes-Coromina et al., 2018). In previous work in which beat gesture and pitch accent were varied orthogonally (Morett & Fraundorf, 2019), no main effect of viewing beat gesture on memory for contrastive information in spoken discourse was observed. Given that this null result was observed in a single study, however, replication is necessary for stronger conclusions concerning beat gesture's (in)ability to serve as a cue to contrast to be drawn.

Beat Gesture and Pitch Accent as Probabilistic Cues

A second question of interest in the present study is how the use of each of these cues is affected by their reliability. For instance, one talker may use beat gesture quite frequently to emphasize contrasts whereas another talker may almost never do so. Thus, interpretation of one cue to prominence may be qualified by whether another cue to prominence is present (or absent) concurrently or used by the talker in general.

Evidence suggests that cue reliability affects how cues, such as beat gesture and pitch accent, are interpreted both alone and conjointly with respect both to the content of co-occurring speech and to the occurrence of other similar cues. Recent work in which beat gesture and pitch accent were manipulated orthogonally with respect to one another indicates that co-occurrence of these cues is particularly effective in enhancing memory (Igalada et al., 2017; Kushch & Prieto, 2016; Kushch et al., 2018; Llanes-Coromina et al., 2018), suggesting that co-occurrence of these cues is interpreted as indicative of importance within this context. This interpretation is consistent with evidence that, in a context in which beat gesture and pitch accent were manipulated orthogonally with respect to one another and focused information, semantic access was impaired by the absence of beat gesture relative to pitch accenting, as evidenced by an increased N400 event-related potential, relative to when both cues co-occurred with focused information (Wang & Chu, 2013). Moreover, it is consistent with evidence that, in a context in which focus was manipulated and beat gesture always co-occurred with pitch accenting, these co-occurring cues disrupted semantic processing when they occurred in conjunction with nonfocused information, as evidenced by an increased late positive potential, relative to when they occurred in conjunction with focused information (Dimitrova et al., 2016). Finally, as we note above, the discrepant findings of Roustan and Dohen (2010a) and Krahmer and Swerts (2007) could potentially be explained by whether or not beat gesture always co-occurred with contrastive accenting. Taken together, these findings *suggest* that the reliability with which pitch accent and beat gesture convey prominence in co-occurring speech influences their effect on spoken language comprehension.

However, we are aware of only one study in which the probability of beat gesture was *directly* manipulated relative to the probability of contrastive accenting with respect to selected alternatives from a contrastive pair (Morett & Fraundorf, 2019). In one discursive context of this study, beat gesture, like contrastive accenting, sometimes accompanied contrastive alternatives; in another context, beat gesture never occurred. In the context in which beat gesture sometimes accompanied contrastive alternatives, contrastive pitch accenting facilitated recognition of contrastive information beat gesture was present but not when beat gesture was absent; by contrast, in the context in which beat gesture never occurred, contrastive pitch accenting reliably facilitated recognition of contrastive information. These findings reveal that the reliability of beat gesture as a cue to contrast influences the extent to which contrastive accenting is interpreted as a cue to contrast and thereby enhances representation of contrastive information. While these findings are provocative in light of evidence that offline memory for discourse is systematically related to online interpretation (Christianson et al., 2001; Huang & Arnold, 2016; Slattery et al., 2013), they do not address whether the effects of these cues emerged at the point of interpretation

or consolidation into memory. The current work addresses this important issue, providing insight into the timing of cue integration during discourse processing.

Specifically, by varying—across experiments—the probability with which beat gesture co-occurs with contrastive information, we investigate whether beat gesture and spoken discourse are integrated via bi-directional *obligatory* interactions regardless of semantic content, as postulated by the integrated systems hypothesis of gesture-speech processing (Kelly et al., 2010), or whether their integration is affected by felicity of beat gesture with contrast within the global linguistic context, such that beat gesture is interpreted as a more valid cue to contrast when it frequently co-occurs with contrastive information. In addressing these questions, the current work provides insight into the broader questions of whether the probability with which multimodal cues, such as gesture and pitch accenting, occur with respect to one another and key information promotes prediction as well as interpretation of linguistic input (see Huettig, 2015, for discussion of this principle) as well as whether previous linguistic input affects cue interpretation during online language processing (Kleinschmidt et al., 2012).

Present Research

Broadly, the purpose of the present research was to investigate how predictions are formulated during online language comprehension based on cue occurrence within discursive contexts. More specifically, we examined whether—and how—beat gesture is interpreted contrastively based on whether it occurs in conjunction with contrastive pitch accenting and contrastive information in local and global discursive contexts. To do so, we used a modified visual world paradigm that incorporated videos of a talker producing beat gesture, similar to paradigms used to examine how iconic gesture is integrated with the semantic content of speech during online reference resolution (Silverman et al., 2010).

In experimental trials of our task, participants heard pairs of context and critical sentences instructing them to click on referents contrasting in color (e.g., *Click on the blue triangle* → *Now, click on the red triangle*) or differing in both color and shape (e.g., *Click on the blue triangle* → *Now, click on the red square*) while we tracked their gaze to objects representing possible referents. In critical sentences, we independently manipulated the presence of beat gesture and contrastive accenting in conjunction with the color word to examine how felicity of these cues with contrastive information within the local referential context, operationalized via contrast type (color-contrast vs. color + shape difference), influenced reference resolution. We predicted that, when only the color word differed between context and critical sentences (i.e., color-contrast; *blue triangle* → *red triangle*), the presence of locally felicitous beat gesture and/or contrastive accenting on the color word would encourage the correct contrastive interpretation (color-contrast), facilitating reference resolution. Furthermore, we predicted that, when both the color and shape words differed between context and critical sentences (i.e., color- + shape-difference, *blue triangle* → *red square*), the presence of locally infelicitous beat gesture and/or contrastive accenting on the color word would misleadingly suggest a referent differing from the context referent in color but not shape (*red triangle*; color-contrast), hindering rather than facilitating reference resolution. By allowing us to investigate the independent and interactive influences of beat

gesture and contrastive accenting on online spoken language processing, the current work differs from previous research examining the influence of contrastive accenting alone on online spoken language processing (Ito et al., 2012; Ito & Speer, 2008; Kurumada et al., 2014; Watson et al., 2008; Weber et al., 2006) and from research investigating the influences of beat gesture and contrastive accenting on offline measures of spoken language processing, such as discourse memory (Kushch & Prieto, 2016; Llanes-Coromina et al., 2018; Morett & Fraundorf, 2019).

A second question of interest was the degree to which interpretation of these cues is guided by their global felicity with contrastive information within a particular linguistic context, as suggested by recent work demonstrating sensitivity to context-specific cue use (Grodner & Sedivy, 2011; Morett & Fraundorf, 2019; Roettger & Franke, 2019; Roettger & Rimland, 2020; Ryskin et al., 2019). We tested this by varying, across experiments, cue contingencies within filler trials to examine how the global felicity of these cues influenced their interpretation on critical trials, as has been done in previous work examining the impact of prosody on scalar implicature interpretation (Huang & Snedeker, 2018). In all cases, filler trials consisted of a pair of sentences instructing participants to click on referents contrasting in shape (e.g., *Click on the blue triangle* → *Now, click on the blue square*) or differing in neither color nor shape (e.g., *Click on the blue triangle* → *Now, click on the blue triangle again*). In filler trials of Experiment 1, beat gesture and contrastive accenting always accompanied contrasting shape words and never accompanied non-contrasting shape words and were therefore globally felicitous; by comparison, in filler trials of Experiment 2, we independently manipulated the presence of beat gesture and contrastive accenting in conjunction with the non-contrasting color word, such that they were globally infelicitous. We predicted that beat gesture and contrastive accenting would encourage contrastive interpretation of critical items when their probability of co-occurring with contrastive information across trials was high (Experiment 1), whereas they would no longer encourage a contrastive interpretation when their probability of co-occurring with contrastive information across trials was low (Experiment 2).

Experiment 1: Felicitous Context

In Experiment 1, we examined the effects of beat gesture and contrastive accenting on online reference resolution in spoken discourse within a linguistic context in which these cues always co-occurred with contrasting words in filler trials, as they typically do within spontaneous natural discourse, and were therefore globally felicitous.

Method

Participants.—Forty adult monolingual native English speakers (age range: 18–35 yrs.; 29 females, 11 males) participated in Experiment 1 and a separate unrelated electrophysiological study in return for \$25 USD. All participants were recruited from the community surrounding a large private research university in the Northeast US via electronic and hard copy announcements. All participants had normal hearing and normal or corrected-to-normal vision and were not colorblind. All participants provided informed consent

to participate, and all experimental procedures were approved by the host university's institutional review board.

Design.—To examine how felicity of cues to contrast with the local referential context affects reference resolution, pitch accenting, beat gesture, and contrast type were varied orthogonally in critical referring expressions, resulting in a 2 (contrastive accenting vs. non-contrastive accenting) x 2 (beat vs. no beat) x 2 (color-contrast vs. color- + shape-difference) design. Eight lists counterbalanced by these independent variables (Lists 1.1–2.4) were used in both practice and experimental trials. Because beat gesture and pitch accenting were manipulated on color words, their presence was locally felicitous in the color-contrast condition and locally infelicitous in the color- + shape-difference condition, whereas the opposite was true for their absence. Figure 1 displays the sentences and accompanying visual displays used in the first critical and filler trials of lists 1.1–1.4. (See Table S1 in Experimental Design Tables, publicly available at <https://osf.io/qx9tf/>, for additional information concerning trials presented in each list.)

Materials.

Objects.—A total of 64 objects (representing all possible combinations of the 8 colors and 8 shapes listed in Table 1) were created for use in arrays accompanying audio and video stimuli during practice and experimental trials. Objects consisted of four types: the context object (i.e., the referent of the context referring expression: e.g., *blue triangle*); the target object (i.e., the referent of the critical referring expression: e.g., *red triangle* [color-contrast] or *red square* [color- + shape-difference]); the competitor object (i.e., an object with the same color but an alternate shape relative to the target object (e.g., *red square* [color-contrast] or *red triangle* [color- + shape-difference]); and the distractor object (i.e., an object with the remaining combination of colors and shapes from other objects in the array: e.g., *blue square*). Critically, the presence of the competitor object created a temporary referential ambiguity (*Now click on the red...*) between the target and competitor objects, allowing us to examine how cues to contrast affected fixations on them when it was unclear which one was the referent.

Across all trials, objects corresponding to all four possible continuations (color-contrast, color- + shape-difference, shape-contrast, neither-difference) were present in the array, and objects corresponding to all four possible continuations were equally likely to follow the context object. Each of the 64 objects was assigned as the context object, target object, competitor object, and distractor object in 2–3 trials total in a variety of conditions in order to avoid contingencies to the greatest extent possible; however, a perfectly equal number of assignments for each individual object was not possible because the number of trials was specified based on the independent variables (contrast type, pitch accent, beat gesture), and counterbalancing of them was prioritized.

During all trials, objects appeared in one of four locations arranged in a square configuration surrounding centrally-presented, circularly-framed videos (see Figure 1). For each item, positioning of objects was counterbalanced across participants such that context, target,

competitor, and distractor objects were equally likely to appear in each position (see Tables S4–S5 in Experimental Design Tables).

Audio recordings.—All sentences audio recorded for use in the experiment were produced by an adult female monolingual speaker of Standard American English. To ensure that audio recordings of sentences were of the highest possible quality, audio recording was conducted using professional equipment in a sound-shielded room prior to video recording.

Each sentence instructed participants to click on a specific object from an array (see Fig. 1 for example array and 7–10d for example sentences). Each trial consisted of a pair of sentences, the first referring to a context object (context sentence; 7, 9) and the second referring to a critical object (continuation sentence; 8a-d, 10a-d; see Supplemental Material for complete list of sentences used in Experiment 1). Within each stimulus list used in the experiment (Lists 1.1–1.4, 2.1–2.4), referring expressions within context and continuation sentences were counterbalanced, such that each critical sentence appeared in trials with all possible contrast types. (See Table S2 in Experimental Design Tables for number of sentences constructed with each pitch accent within each list set).

Lists 1.1–1.4

- (7) [Context] Click on the blue star.
- (8a) [Color-contrast (critical)] Now click on the white star.
- (8b) [Color- + shape-difference (critical)] Now click on the white square.
- (8c) [Shape-contrast (filler)] Now click on the blue square.
- (8d) [Neither-difference (filler)] Now click on the blue star again.

Lists 2.1–2.4

- (9) [Context] Click on the blue square.
- (10a) [Color-contrast (critical)] Now click on the white square.
- (10b) [Color- + shape-difference (critical)] Now click on the white star.
- (10c) [Shape-contrast (filler)] Now click on the blue star.
- (10d) [Neither-difference (filler)] Now click on the blue square again.

During audio recording, the talker pictured the same array while recording both a context and a continuation sentence. The sentences were constructed so that any combination of colors and shapes represented in the array could serve as the target in the continuation sentence (see Table 1).

For each critical continuation sentence, we created one recording in the *contrastive accenting* condition and one in the *non-contrastive accenting* condition. In the contrastive accenting condition, the talker was instructed to emphasize the color adjective and produced the context sentence followed by the corresponding color-contrast continuation sentence; in the non-contrastive accenting condition, the talker was instructed not to emphasize any

particular word and produced the context sentence followed by the corresponding color- and shape-difference continuation sentence. (See Supplemental Material for audio recording list.) Acoustic analyses, reported below, confirmed that, as expected, the contrastive accenting condition resulted in greater acoustic prominence (Selkirk, 2002), supporting the validity of our manipulation.

The entire list of sentences was recorded twice in two separate sessions. One recording of each critical sentence with non-contrastive (presentational) accenting was selected for use of its color word (non-contrastive accenting sentence), and the other recording was selected for use of its other components (carrier sentence). In addition, one recording of each critical sentence with contrastive accenting was selected for use of its color word (contrastive accenting sentence). Finally, one recording of each filler sentence with non-contrastive and contrastive accenting was selected for wholesale use. When possible, recordings were selected on the basis of their quality and the prototypicality of their pitch accents. In post-production, all sentences were labeled and batch exported as individual wav files using Audacity (Audacity Team, 2013). Audio files of final sentences were copied as needed and used in all trials in which they were required based on stimulus lists and the experimental design.

Critical sentences.—To allow for manipulation of cue felicity with the local referential context, the target object of critical experimental trials corresponded to one of two contrast types: Relative to the preceding (context) object, half differed in color but not in shape (color-contrast; 8a, 10a), and half differed in both color and shape (color- + shape-difference; 8b, 10b). In critical sentences in these trials, half of the color words received a contrastive (L+H*) accent and the other half of the color words received a non-contrastive (presentational; H*) pitch accent. Thus, contrastive accenting of color words was locally felicitous with color-contrast but not color- + shape-difference referring expressions, whereas non-contrastive accenting of color words was locally felicitous with color- + shape-difference but not color-contrast referring expressions. All shape words in these sentences were naturally de-accented.

Using Praat (Boersma & Weenink, 2016), text grids for sentences with non-contrastive and contrastive accenting and carrier sentences were created in which initial portions (*Now click on the...*), color words (*purple*), stressed syllables of multisyllabic color words (*purp-*), and shape words (*triangle*) were annotated. These text grids were subsequently used to batch extract these components as individual files. For color words and stressed syllables of multisyllabic color words in sentences with non-contrastive and contrastive accenting, intensity, duration, maximum F_0 (fundamental frequency), differences between maximum and minimum F_0 , and mean F_0 were measured and subsequently analyzed using R.

To verify that color words used in critical sentences indeed differed between the contrastive and non-contrastive pitch accent conditions, we compared their acoustic properties using a series of paired-sample *t*-tests. Because it is argued that pitch accenting is realized on the syllable carrying the primary stress of a word (Ladd, 1996), analyses were conducted on the stressed syllable of multisyllabic color words as well as entire color words. Table 2 presents means and standard deviations of acoustic measures. Both stressed syllables and color words

with contrastive and non-contrastive pitch accenting differed reliably on all measures except F_0 difference.

To eliminate any spurious differences in intensity across individual recordings, intensity was normalized to the average across items of each component (initial, color word, shape word) for each critical sentence type (accenting, carrier). Finally, color words from non-contrastive and contrastive accenting sentences were spliced between initial portions and shape words of carrier sentences. This ensured that the only part of the sentence that differed acoustically or phonologically between the two types of critical referring expressions presented in these trials was the color word. To ensure that splicing of critical sentences was not noticeable, a subset of spliced and unspliced critical sentences was played to a pilot participant, who was unable to reliably identify which sentences were spliced.

Filler sentences.—Filler trials were included to ensure that all types of referents sometimes served as the target in the continuation referring expression, thus preventing prediction of the target merely based on knowledge of which types of objects could be referred to within the experimental task. Thus, half of filler referring expressions differed in shape but not in color relative to the preceding context referring expression (shape-contrast; 8c, 10c), and half differed in neither shape nor color (neither-difference; 8d, 10d). In Experiment 1, shape words always had contrastive accenting in shape-contrast referring expressions and non-contrastive accenting in neither-difference referring expressions, such that pitch accenting was globally felicitous. This differed from Experiment 2, in which, similar to critical referring expressions, pitch accenting was orthogonally manipulated on the *color* word in shape-contrast and neither-difference referring expressions, such that it was globally infelicitous. (See Table S3 in Experimental Design Tables for summary of filler item attributes.)

Filler sentences were annotated, segmented, and analyzed for acoustic properties similarly to critical sentences. For fillers, we expected the primary accent to fall on the shape word (e.g., *triangle*) because the fillers either involved a shape contrast or no contrast, in which the pitch accent would be expected to fall on the head shape noun (e.g., Selkirk, 1995). Thus, acoustic analyses were based on shape words (e.g., *triangle*) and stressed syllables of multisyllabic shape words (e.g., *tri-*). To verify that the contrastive and non-contrastive accenting conditions differed acoustically, we compared them using a series of paired-sample *t*-tests. Table 2 presents means and standard deviations of acoustic measures. Stressed syllables with contrastive and non-contrastive pitch accenting differed reliably on all measures except maximum F_0 and F_0 difference, whereas shape words with contrastive and non-contrastive pitch accenting differed reliably on all measures except intensity.

Again, to eliminate any differences in intensity across individual recordings, intensity was normalized to the average for filler sentences with each type of pitch accent (non-contrastive, contrastive). Unlike critical sentences, filler sentences were not spliced. Pupillometric and mouse click responses to filler sentences were not analyzed because only the critical trials (by definition) included the parametric manipulations of beat gesture and contrastive accenting.

Video recordings.—Videos to accompany all sentences presented in the experiment were recorded separately from audio recordings. Videos featured a Caucasian adult female (henceforth, the talker) and were framed on her torso with her head excluded to ensure that facial cues could not affect sentence interpretation. To demonstrate to the talker how beat gesture should be produced during video recording of sentence production, the first author (L.M.M.) modeled one of the most common beat gestures that talkers produce in natural conversation: a rapid single downward stroke of the dominant (right) hand¹ with the palm open upward (McNeill, 1992). During video recording, the talker listened to audio recordings of sentences and subsequently repeated them while either producing a beat gesture in conjunction with the color or shape word or while keeping her hands still (see Table S6 in Experimental Design Tables for number of videos constructed for each set of lists). The talker was instructed never to produce beat gesture while producing context sentences (7, 9).

Prior to post-production, we reviewed the videos and confirmed that the talker produced the correct pitch accent in the appropriate place² and that the talker's beat gesture and pitch accenting were temporally aligned. In post-production, we used Adobe Premiere Pro to temporally align videos with audio recordings, such that beat gesture stroke onsets occurred 200 ms prior to color word onsets, resulting in apices co-occurring with stressed syllables of color words. This ensured that the apex of the beat gesture's stroke occurred 200 ms prior to the onset of the corresponding word, consistent with the timing of gesture production relative to natural spoken discourse (Morrel-Samuels & Krauss, 1992) and with perceptual biases for gesture relative to speech (Leonard & Cummins, 2011). Videos were then trimmed to the length of audio recordings. A white circular reverse mask was then added to videos so that they appeared as circular against the white background used in the experimental paradigm, ensuring that arrays of objects were equidistant from centrally-located, circularly-framed videos in visual displays (similar to the implementation of video in the visual world display used in Silverman et al., 2011; see Figure 1).

Critical sentences.—To allow for manipulation of felicity of beat gesture with the local referential context, two videos of each critical sentence (8a-b, 10a-b) were recorded consecutively: one in which the talker produced a beat gesture in conjunction with the color word, and one in which the talker did not produce beat gesture. Thus, beat gesture occurred only during the color word within color- + shape-difference critical referring expressions, ensuring that its presence and timing were identical during color-contrast and color- + shape-difference sentences. Although it is possible that beat gesture produced in conjunction with contrastive accenting may have differed qualitatively from beat gesture produced in conjunction with non-contrastive accenting, recording separate videos for critical sentences with contrastive and non-contrastive accenting allowed us to optimize gesture-accent duration synchrony and ecological validity while also permitting us to

¹To control for possible associations between the side on which the beat gesture occurred and the side on which the target object appeared, horizontally-flipped duplicates were created for all videos, and presentation of original and flipped videos was counterbalanced across trials and participants by list (see Tables S4–S5).

²Because the original soundtracks of video clips were deleted, it was not possible to analyze the acoustic features of pitch accented words as was done for the audio-recorded spliced sentences that replaced them.

manipulate beat gesture orthogonally relative to pitch accenting in these trials, crucial to probing interactions between these two cues to contrast.

Filler sentences.—Because beat gesture and pitch accent were not manipulated orthogonally in filler sentences as they were in critical sentences, a single video was recorded to accompany each filler sentence (8c-d, 10c-d). In videos accompanying shape-contrast sentences, the talker produced a beat gesture in conjunction with the shape word, and in videos accompanying neither-contrast sentences, the talker did not produce a beat gesture. This configuration maintained the pattern of usage of, and association between, pitch accenting and beat gesture found in natural spoken discourse, such that these cues were globally felicitous.

Norming.—We conducted norming to confirm that the co-occurrence of beat gesture and contrastive accenting and their local felicity with contrast in critical and filler sentences were perceived as expected. 78 participants, who did not participate in either Experiment 1 or 2, watched a sample of stimuli from Experiments 1 and 2 and rated each of them on a 1 (completely unnatural) – 7 (completely natural) basis for (a) gesture-accent co-occurrence (*How well did the speaker's gestures match the speech?*) and (b) local felicity of these cues with contrast (*How well did the speech style fit the instructions given?*). The co-occurrence of beat gesture and contrastive accenting was rated as more natural than either cue in isolation. Further, both cues were perceived as more natural when they were locally felicitous with contrast. These findings confirm that beat gesture and contrastive accenting are perceived as most natural when they co-occur with one another and are locally felicitous with contrast. The full results of this norming study can be found in the Supplementary Material section.

Procedure.—Prior to beginning the experimental task, participants were seated 55–65 cm from the monitor on which stimuli were presented ($35^{\circ} 55' 0.32''$ visual angle). Gaze was calibrated to within 0.5° of visual angle using 13 points of reference. Drift checks were performed between each of the experimental trial blocks, and recalibration was performed if gaze was misaligned by more than 2° of visual angle.

At the beginning of the experimental task, participants were told that its purpose was to test their ability to follow instructions. Consequently, they were instructed to respond to all instructions as quickly and accurately as possible by using a mouse to click on the appropriate shape. They were also instructed that, if they accidentally clicked on the wrong shape, they would need to click on the correct shape to proceed; however, in both experiments, all responses to critical continuation referring expressions were correct. No instructions concerning gaze were provided.

Participants first completed an 8-trial practice block to become familiar with the task, then proceeded to the experimental blocks (4 blocks of 40 trials each; see Table S8 in Experimental Design Tables). In both practice and experimental blocks, critical and filler trials were randomly interleaved. Each trial began with a context sentence accompanied by a visual display consisting of a centrally-located video and a surrounding array of objects corresponding to that sentence and the subsequent continuation sentence (see Figure 1).

Following a correct response, the video disappeared and was replaced by a gray circular placeholder for 1000 ms while the object array remained on screen to emphasize the continuity between context and critical sentences. Subsequently, the sequence repeated with the continuation sentence and its corresponding video. Following a correct response, the trial ended and, following a blank screen displayed for 1000 ms, a new trial began or the block ended. After each block, participants received the opportunity to take a brief break. Once they indicated that they were ready to begin the next block, a drift check was performed and, if necessary, gaze was recalibrated prior to the first trial of that block.

During experimental trials, participants were permitted to gaze freely at the monitor while gaze data were collected remotely from the right eye at a 500 Hz sampling rate using an EyeLink 1000 eye-tracker (SR Research). If participants' gaze left the trackable range, an audible alert sounded and participants were re-positioned and re-calibrated if necessary.³

Results

Our interest was in how the temporary referential ambiguity created by critical continuation sentences (e.g., *Now click on the red...*) was resolved. Thus, our predictions concerned the frequency with which target objects (the eventual referents of critical referring expressions) and competitor objects (referents temporarily consistent with unfolding critical referring expressions but ultimately incorrect) were fixated.⁴ Because all mouse click responses to critical sentences were correct, we analyzed all eye gaze data collected during critical sentence processing.

We examined these fixations within each of three time periods of interest. First, we examined the *trial onset interest period*, which lasted from the onset of the trial to the onset of the color word; because this period terminates before any referential expression, it tests for fixation biases to objects prior to disambiguating input. Next, we examined the *color word interest period*, which lasted from the onset of the color word to the onset of the shape word to determine whether it was possible to anticipate the correct referent on the basis of beat gesture and/or pitch accent during the temporarily ambiguous referring expression. The durations of both trial onset and color word interest periods were determined on a trial-by-trial basis; however, these interest periods are represented in figures via average duration across trials. Finally, we examined the *shape word interest period*, which lasted from the onset of the shape word to the point at which fixations peaked across trials (930 ms after shape word onset) to evaluate whether beat gesture and/or pitch accenting affected reference resolution following disambiguating lexical input (but prior to execution of any extraneous fixations following target identification). To account for fixation planning, 200 ms was added to the onsets of color and shape words at the beginning and end of interest periods.

³This occurred in <1% of trials.

⁴In Expts. 1 and 2, we opted to analyze target and competitor fixations separately rather than analyzing the ratio of target to competitor fixations because there was greater competition between target and distractor objects than between target and competitor objects in Expt. 2 (see Fig. 5), and we wanted to use the same dependent variables in both experiments to facilitate comparison between their results.

Prior to entry into models, fixation data was converted into empirical logit values by aggregating over samples within a trial (Barr, 2008).⁵ All models included fixed effects of beat gesture (beat, no beat) and pitch accent (contrastive, non-contrastive) to examine how these cues to contrast affected reference resolution, contrast type (color, color + shape) to examine differential effects of these cues on resolution of referents with a specific contrast vs. a general difference, and trial to examine how these effects changed over time, as well as interactions between these factors. In addition, all models included fixed effects of gesture orientation (left, right) and object side (left, right) and their interaction to examine whether congruency between the side on which beat gesture and the target object occurred affected fixations. All fixed effects were coded using mean centered (Helmert) contrast coding, with the level mentioned first for each factor coded as -0.5 and the level mentioned second coded as $+0.5$. In all models, the maximal random effects structure permitting convergence was used to minimize Type I error. All models were fit in R using the `lmer()` function of the *lme4* package (Bates et al., 2015). Null hypothesis significance testing was conducted using the *lmerTest* package (Kuznetsova et al., 2017). For interactions reaching significance, Tukey HSD post-hoc tests were conducted using the *emmeans* package (Lenth, 2019), and comparisons of the effects of beat gesture and pitch accenting within each contrast type are reported where appropriate. All data and analysis scripts are publicly available via the following link: <https://osf.io/qx9tf/>.

Video Fixations.—One concern is that apparent differences across conditions in fixations to the target object might be driven by fixations to the video, especially in the conditions in which the video contains a beat gesture. To address this concern, we examined, for each interest period, fixations to the video interest area vs. all object interest areas. As can be seen in Table 3, during the trial onset and color word interest periods, participants fixated the video far more than all other interest areas. During the shape word interest period, however, fixations to the video decreased substantially whereas fixations to the target and competitor objects increased substantially, indicating that participants tended to look at those objects when resolving referring expressions. To determine whether participants were more likely to fixate the video during each interest period when beat gesture was present than when it was absent, we modeled the empirical logit of fixations to the video vs. all other interest areas during each interest period using the effect structure specified above. These analyses failed to reveal a main effect of beat gesture on fixations to the video in any interest period (see Table 3 for proportions and Tables 4–6 for parameter estimations of fixations on video during each interest period). Because there was no evidence of a difference in video fixations across conditions, and our primary interest was in the effect of these conditions on resolution of the temporary ambiguity among potential referents (rather than referents versus a video), video fixations were excluded from the main analyses.

Overall Time Course.—Figure 2 displays the proportion of fixations on each type of object over time during critical referring expressions. This figure suggests, as expected, that competition between the target and competitor (both temporarily consistent with the unfolding linguistic input) was eventually correctly resolved in favor of the target.

⁵Descriptive statistics and plots are based on raw proportions of fixations rather than the empirical logit to facilitate interpretation.

The critical question was whether—and how—beat gesture and pitch accenting affected resolution of this ambiguity differently by contrast type in critical items within a linguistic context in which these cues were globally felicitous with contrastive information in filler items. Specifically, we sought to determine whether the presence of these cues during the color word of color-contrast referring expressions increased fixations to the target and decreased fixations to the competitor during the color and shape word interest periods. Conversely, we sought to determine whether the presence of these cues during the color word of color- + shape-contrast referring expressions decreased fixations to the target and increased fixations to the competitor during the color and shape word interest periods. Figure 3 displays the proportion of fixations on targets and competitors by pitch accenting, beat gesture, and contrast type over time during critical sentences, which we submitted to separate models for each interest period.

Trial Onset Interest Period.

Target fixations.: We observed a significant main effect of beat gesture, indicating that more target fixations occurred slightly more during trial onset when beat gesture was absent ($M = 0.022$, $SD = 0.157$) than when beat gesture was present ($M = 0.021$, $SD = 0.146$; see Table 7). This difference, which is barely discernible in the trial onset IP of the left two panels of Figure 3, may have resulted from greater attendance to the video during trial onset due to preparatory hand movement in trials with beat gesture.

Competitor fixations.: No main effects or interactions of any of the factors of interest were observed (see Table 8).

Color Word Interest Period.

Target fixations.: A significant interaction between beat gesture and contrast type was observed for target fixations during the color word (see Table 9). To further explore this interaction, we conducted Tukey HSD-corrected post-hoc tests (see Table 10). During color-contrast referring expressions, there was no significant difference in target fixations during the color word when beat gesture was present ($M = 0.037$; $SD = 0.202$) compared to when beat gesture was absent ($M = 0.036$; $SD = 0.187$), as is evident in the overlap between the lines within the color word IP of the top left panel of Figure 3. During color- + shape-difference referring expressions, however, fewer target fixations occurred during the color word when beat gesture was present ($M = 0.026$; $SD = 0.159$) compared to when beat gesture was absent ($M = 0.035$; $SD = 0.190$), as is evident in the gaps between the respective lines within the color word IP of the bottom left panel of Figure 3. That is, when critical referring expressions differed in both color and shape from a previous context referring expression, emphasizing the color word with beat gesture reduced anticipatory fixations towards its referent. This interaction was qualified by a differently-signed interaction with trial, indicating that this effect decreased over the course of the experiment.

Competitor fixations.: An interaction between beat gesture and contrast type was also observed for competitor fixations during the color word (see Table 11). To further explore this interaction, we conducted Tukey HSD-corrected post-hoc tests (see Table 12). Examination of cell means revealed that, during the color word of color-contrast referring

expressions, there was no significant difference in fixations on color- + shape-difference competitors when beat gesture was present ($M = 0.024$; $SD = 0.152$) compared to when beat gesture was absent ($M = 0.024$, $SD = 0.168$), as is evident in the overlap between the lines within the color word IP of the top right panel of Figure 3. During the color word of color- + shape-difference referring expressions, however, more fixations on color-contrast competitors occurred when beat gesture was present ($M = 0.052$, $SD = 0.024$) compared to when beat gesture was absent ($M = 0.034$; $SD = 0.187$), as is evident in the gaps between the respective lines within the color word IP of the bottom right panel of Figure 3. That is, when a critical referring expression differed in both color and shape from the previous context referring expression, the presence of beat gesture during a color word misled comprehenders into anticipating the color-contrast competitor.

Shape Word Interest Period.

Target fixations.: We observed a main effect of contrast type indicating that more target fixations occurred during the shape word of critical referring expressions that contrasted with context referring expressions only in color ($M = 1.088$; $SD = 0.739$) than that differed from context referring expressions in both color and shape ($M = 0.933$; $SD = 0.705$; see Table 13), as is evident in the steeper slope of all lines in the top left panel compared to the top right panel of Figure 3. We also observed a negative main effect of trial indicating that target fixations decreased over the course of the experiment. These two main effects interacted, indicating that the difference in target fixations during color-contrast and color- + shape-difference critical referring expressions decreased over the course of the experiment.

We also observed significant three-way interactions between contrastive accenting, beat gesture, and trial and between beat gesture, contrast type, and trial. Because the interactions between contrastive accenting and beat gesture and between beat gesture and contrast type failed to reach significance, they did not affect overall target fixations, however. Moreover, because the three-way interactions with trial were differently-signed than the two-way interactions between contrastive accenting and beat gesture and between beat gesture and contrast type, differences in target fixations resulting from these interactions decreased in magnitude over the course of the experiment.

In addition to the effects of the main factors of interest, the interaction between gesture orientation and object side reached significance for target fixations. This result suggests that targets appearing on the side of the array congruent with the orientation of beat gestures were fixated more than targets appearing on the side of the array incongruent with the orientation of beat gestures during shape words in critical referring expressions. Because of counterbalancing, however, this effect was orthogonal to our variables of interest, so we do not discuss it further.

Competitor fixations.: We also observed a positive main effect of trial indicating that competitor fixations increased over the course of the experiment. We also observed a main effect of contrast type indicating that more competitor fixations occurred when critical referring expressions differed from context referring expressions in both color and shape ($M = 0.428$; $SD = 0.602$) than when they contrasted only in color ($M = 0.245$; $SD = 0.499$),

as is evident in the steeper curves in all lines within the shape word IP of the bottom right compared to the top right panel of Figure 3. This main effect was qualified by an interaction with contrastive accenting (see Table 14). To further explore this interaction, we conducted Tukey HSD-corrected post-hoc tests (see Table 15). Examination of cell means revealed that, during the shape word of color- + shape-difference referring expressions, more fixations occurred on color-contrast competitors when the color word was contrastively accented ($M = 0.448$; $SD = 0.592$) than when the color word was non-contrastively accented ($M = 0.409$, $SD = 0.612$), as is evident in the steeper curves of the lines for conditions with contrastive accenting than the lines for conditions without contrastive accenting within the shape word IP of the bottom right panel of Figure 3. During the shape word of color-contrast referring expressions, however, there was no significant difference in fixations on color- + shape-difference competitors when the color word was contrastively accented ($M = 0.263$; $SD = 0.521$) compared to when the color word was non-contrastively accented ($M = 0.228$, $SD = 0.475$), as is evident in the overlap between the lines within the shape word IP of the top right panel. These results provide further evidence that interpretation of critical referring expressions that did not specifically contrast in color with the previous referent was hindered by contrastive accenting on the color word, which continued to mislead comprehenders during reference resolution by encouraging them to continue considering the color-contrast competitor as a possible referent.

In addition to the effects of the main factors of interest, the interaction between gesture orientation and object side reached significance for competitor fixations. This result suggests that competitors appearing on the side of the array congruent with the orientation of beat gestures were fixated more than competitors appearing on the side of the array incongruent with the orientation of beat gestures during shape words in critical referring expressions. Because of counterbalancing, however, this effect was orthogonal to our variables of interest, so we do not discuss it further.

Discussion

We examined how beat gesture and contrastive accenting affect online processing of reference to objects that either contrasted in color with the prior context referent, such that the presence of these cues on the color word was felicitous with the local referential context, or differed in both color and shape from the prior context referent, such that the presence of these cues on the color word was infelicitous with the local referential context. In Experiment 1, we examined these effects in a global linguistic context in which beat gesture and contrastive accenting were always felicitous with contrastive information in filler items, such that they served as highly reliable cues to contrast across trials. Although the results failed to support our prediction that these cues would increase fixations to the color-contrast target and decrease fixations to the color- + shape-difference competitor of color-contrast referring expressions, they supported our prediction that these cues would decrease fixations to the color- + shape-difference target and increase fixations to the color-contrast competitor of color- + shape-difference referring expressions. Thus, they provide evidence that beat gesture and contrastive accenting are interpreted contrastively within linguistic contexts in which they reliably convey contrast.

Relative to previous eye-tracking research using visual world paradigms to examine contrastive accent processing (Ito et al., 2012; Ito & Speer, 2008; Kurumada et al., 2014; Watson et al., 2008; Weber et al., 2006), convergence of fixations to the target object was slower and the maximum proportion of fixations to target and competitor objects was lower across conditions. Both of these effects are likely due to the high proportion of fixations to the video relative to objects during the entirety of the critical sentence, particularly during the trial onset and color word interest periods. Indeed, no previous visual world eye-tracking research examining contrastive accent processing has included videos, and the only other visual world eye-tracking study that we are aware of that included videos showed proportions of fixations to target and competitor interest areas for healthy young adults that were at least 500 ms later and were halved in proportion relative to comparable visual world eye-tracking research that did not include videos (Silverman et al., 2010). Thus, the results of this previous study and the current work indicate that inclusion of visual linguistic cues via video in visual world paradigms substantially decreases the proportion and speed of fixations to target and competitor referents.

Our primary interest was in the effect of the prominence cues on reference resolution. We found that locally-infelicitous contrastive accenting on color words hindered resolution (though not anticipation) of color- + shape-difference referring expressions by encouraging consideration of a competing color-contrast referent. This accords with prior findings from similar paradigms (e.g., Ito & Speer, 2008) in which contrastive accenting is interpreted as an indicator of referential contrast, increasing fixations to competitor objects when it is infelicitous with contrast. Unlike previous research (Ito et al., 2012; Ito & Speer, 2008; Kurumada et al., 2014; Watson et al., 2008; Weber et al., 2006), however, locally-felicitous contrastive accenting failed to facilitate anticipation and resolution of color-contrast referring expressions. Although the reason for contrastive accenting's failure to facilitate identification of contrastive target referents is unclear, the presence of beat gesture may have affected it, given that previous work has demonstrated that beat gesture is interpreted as a more salient cue in linguistic contexts in which both cues occur and are manipulated relative to one another (Kushch et al., 2018; Kushch & Prieto, 2016; Llanes-Coromina et al., 2018; Morett & Fraundorf, 2019).

Our novel finding was that beat gesture also functioned as a cue to contrast: For referents that differed in both color and shape from a previous referent, emphasizing the color word specifically with a beat gesture impaired comprehension by reducing anticipation of the target referent and increasing (incorrect) anticipation of a color-contrast competitor. Reduced anticipation of color- + shape-difference target referents was modulated by trial, however, indicating that comprehenders learned that beat gesture was not always locally felicitous with color contrasts within the context of critical referring expressions and adapted their reference resolution accordingly. We note that, while locally infelicitous beat gesture impaired comprehension of color- + shape-difference referring expressions, its converse *absence* did not impair comprehension of color-contrast referring expressions. That is, using beat gesture to emphasize the color word when there was not a color contrast particularly impaired comprehension, but comprehension was not impaired when color-contrast referring expressions were unaccompanied by beat gestures. This pattern replicates what has sometimes been observed with contrastive accenting (e.g., Watson

et al., 2008; c.f., Ito & Speer, 2008) and suggests that the presence of a prominence cue is a more marked case than its absence. Like contrastive accenting, however, beat gesture failed to enhance anticipation of referents of color-contrast referring expressions when it accompanied the color word both by itself and in conjunction with contrastive accenting. Again, the reason for this finding is unclear, particularly in light of previous findings that beat gesture enhances memory for contrastive information in spoken discourse, particularly when it occurs conjointly with contrastive accenting (Kushch & Prieto, 2016; Llanes-Coromina et al., 2018; Morett & Fraundorf, 2019). Given that the current work examined how beat gesture and contrastive accenting affect online contrast resolution rather than offline memory for selected alternatives from contrastive pairs, one possible explanation is that beat gesture may affect processing of contrastive information differently than it affects memory for contrastive information. Future research should further explore this possibility by directly comparing beat gesture's effects on processing and memory of the same contrastive information in spoken discourse.

Experiment 1 suggests that beat gesture is interpreted as a cue to contrast under relatively favorable conditions in which it always matched the referential context of filler items and was therefore globally felicitous. However, beat gesture may be less likely to be interpreted as a cue to contrast under relatively unfavorable conditions in which it never matches the referential context of filler items and is therefore globally infelicitous. Thus, we conducted Experiment 2 to investigate this possibility. In filler referring expressions of Experiment 2, we manipulated beat gesture and pitch accent orthogonally in relation to the color word, such that these cues never co-occurred with contrastive information, rather than conjointly in relation to the shape word, such that they always co-occurred with contrastive information. Given that, in natural discourse, beat gesture and contrastive accenting would typically occur in conjunction with contrasting shape words rather than non-contrasting color words in these referring expressions, the filler referring expressions of Experiment 1 provided a more ecologically valid context in which these cues reliably indicated contrast, whereas those of Experiment 2 provided a less ecologically valid context in which these cues failed to reliably indicate contrast.

Experiment 2: Infelicitous Context

In Experiment 2, we again examined the effects of beat gesture and contrastive accenting on online reference processing in spoken discourse comprehension, but now within a linguistic context in which these cues were globally infelicitous, permitting us to investigate how the reliability with which these cues convey contrast affects their interpretation.

Methods

Participants.—Forty adult monolingual native English speakers (age range: 18–26 yrs.; 36 females, 4 males) separate from those who participated in Experiment 1 participated in Experiment 2 for partial course credit. All participants were recruited from a large public research university in the Southeastern US via relevant courses. All participants had normal hearing and normal or corrected-to-normal vision and were not colorblind.

Design.—As in Experiment 1, to examine how felicity of cues to contrast with the local referential context affects reference resolution, orthogonal manipulation of pitch accent, beat gesture, and contrast type in critical referring expressions resulted in a 2 (color-contrast vs. color- + shape-difference) x 2 (contrastive accenting vs. no contrastive accenting) x 2 (beat vs. no beat) within-participants design. Once again, eight lists counterbalanced by contrast type, pitch accent, and beat gesture (Lists 1.1–2.4) were used in both practice trials and critical and filler trials. As in Experiment 1, because beat gesture and pitch accenting were manipulated on color words, their presence was locally felicitous in the color-contrast condition and locally infelicitous in the color- + shape-difference condition, whereas the opposite was true for their absence. Figure 4 displays the sentences and accompanying visual displays used in the first critical and filler trials of lists 1.1–1.4. (See Table S6 in Experimental Design Tables for additional information concerning trials presented in each list.)

Materials.

Objects.—The same 64 objects used in Experiment 1 were used in Experiment 2. Objects presented during critical referring expressions were classified according to the same types and assigned in the same way as in Experiment 1, allowing us to examine how cues to contrast affected fixations on them when it was unclear which one was the referent. As in Experiment 1, during all trials, objects appeared in one of four locations equidistant from centrally-presented, circularly-framed videos (see Figure 4). Again, for each item, the positioning of objects was counterbalanced across participants such that context, target, competitor, and distractor objects were equally likely to appear in each position (see Tables S4 and S7 in Experimental Design Tables).

Audio recordings.—The same spliced audio recordings of critical sentences created for Experiment 1 were used as critical sentences in Experiment 2. To prevent participants from anticipating when and how contrastive accenting would occur based on the contrast type of continuation sentences and increase competition between possible referents, new filler sentences were created in which pitch accenting was manipulated orthogonally on the color word, as it was in critical sentences. To create the new filler sentences, we spliced color words with contrastive and non-contrastive accenting from critical sentences into carrier sentences that we created from filler sentences in Experiment 1 with the same combination of color and shape words as new filler sentences.⁶ Although specific filler sentences differed between Experiments 1 and 2 due to this splicing, the number of sentences of each type and accent in the practice and experimental blocks for each set of lists was the same in both experiments.

As in Experiment 1, half of filler referring expressions differed in shape word but not in color word relative to the preceding context (shape-contrast), and half differed in neither shape nor color word (neither-difference). However, in Experiment 2 fillers, pitch accenting differed in *color* words. Because the targets of fillers were always the same

⁶For neither-contrast filler referring expressions, filler sentences from Expt. 1 were used as carrier sentences to create filler sentences for Expt. 2 that preserved the color word + “again” combination, which sounded more natural than splicing color words and “again” together.

color as the preceding context object, the presence of contrastive accenting on color words was always infelicitous, such that pitch accenting was globally infelicitous (see Table S5 in Experimental Design Tables). In order to distribute all pitch accent combinations evenly across sentences with all contrast types, context and critical sentences were paired differently and trials were ordered differently in critical and filler trials than they were in Experiment 1 (see Supplemental Material for complete list of sentences used in Experiment 2).

Video recordings.—The same videos used with critical sentences in Experiment 1 were used with critical sentences in Experiment 2. To construct additional videos to accompany the new filler sentences of Experiment 2 to manipulate beat gesture orthogonally with respect to pitch accent on the color word, we temporally re-aligned video recordings with and without beat gesture used with Experiment 1 fillers to match audio recordings used in Experiment 2 fillers with each pitch accent. (See Table S8 in Experimental Design Tables for number of videos constructed for each set of lists.) Videos were temporally re-aligned with audio recordings of filler sentences such that stroke onsets in videos with beat gesture (temporally re-aligned in Expt. 1 to occur 200 ms prior to shape word onsets in fillers) occurred 200 ms prior to color word onsets, resulting in apices co-occurring with stressed syllables of color words. Videos without beat gesture were similarly temporally re-aligned with audio recordings of filler sentences, such that shape words in audio tracks of videos (prior to removal) co-occurred with color words in audio recordings. Although this configuration maintained the temporal association between pitch accenting and beat gesture, it broke the temporal association between these cues and contrasting information (shape words) found in natural spoken discourse, such that these cues were globally infelicitous. Temporal re-alignment resulted in videos shorter in length than the corresponding audio recordings for Experiment 2 fillers. To fill this gap in timing, the last frame of each video was extended over its duration, resulting in a freeze frame at the end of each video for fillers.⁷

Procedure.

The same procedure used in Experiment 1 was used in Experiment 2.

Results

Fixation data was analyzed using the same procedures described for Experiment 1. All data and analysis scripts are publicly available via the following link: <https://osf.io/qx9tf/>. Because paradigms and participants differed between Experiments 1 and 2, data from each experiment was analyzed and is reported separately here.

Video Fixations.

As can be seen in Table 16, during the trial onset and color word interest periods, participants fixated the video far more than all other interest areas. During the shape word interest period, however, fixations to the video decreased substantially whereas

⁷Because no motion occurred during this final video portion, extending the last frame looked relatively similar to the original videos, although the freeze frame was noticeable.

fixations to the target, competitor, and distractor objects increased substantially, suggesting that participants tended to look at those objects when resolving referring expressions. To determine whether fixations on the video were more likely when beat gesture was present than when it was absent, we modeled the empirical logit of fixations to the video vs. all other interest areas during each interest period using the effect structure specified in Experiment 1. These analyses failed to reveal a main effect of beat gesture in any interest period (see Table 16 for proportions and Tables 17–19 for parameter estimations of fixations on video during each interest period), so we again excluded video fixations from the main analyses reported below.

Overall Time Course.

Figure 5 displays the proportion of fixations on each type of object over time during critical referring expressions. This figure suggests, as expected, that competition between the target, competitor, and distractor (temporarily consistent with the unfolding linguistic input and/or contrast cues) was eventually correctly resolved in favor of the target.

As in Experiment 1, we were interested in whether—and how—beat gesture and pitch accenting affected resolution of this ambiguity differently by contrast type. Our critical question of interest, however, was how global infelicity of beat gesture and contrastive accenting in filler items of Experiment 2 affected resolution of this ambiguity in critical items. Specifically, we sought to determine whether the increases in fixations to the color-contrast competitor of color- + shape-difference critical referring expressions elicited by beat gesture during the color word interest period and contrastive accenting during the shape word interest period in Experiment 1 would decrease in magnitude or disappear. Figure 6 displays fixations on target and competitor objects by contrast type, pitch accenting, and beat gesture across trials over time during critical sentences, which we submitted to separate models for each interest period.

Trial Onset Interest Period.

Target fixations.—We observed a negative main effect of trial, indicating that fixations on the target decreased over the course of the experiment (see Table 20). This main effect was qualified by an interaction of beat gesture and contrast type, indicating that differences in target fixations by contrast type as a function of beat gesture decreased over the course of the experiment. The interaction between beat gesture and contrast type failed to reach significance, however, indicating that any such differences in target fixations were not detectable overall, suggesting that this interaction was driven by the main effect of trial.

Competitor fixations.—Again, we observed a negative main effect of trial, indicating that fixations on the competitor during the trial onset interest period decreased over the course of the experiment (see Table 21). In addition, we observed a main effect of gesture orientation, indicating that fixations on the competitor during the trial onset interest period differed based on the orientation of beat gestures. This main effect was qualified by object side, suggesting that fixations on competitors may have occurred more when they appeared on the side of the array congruent with the orientation of beat gestures than when they appeared on the side of the array incongruent with the orientation of beat gestures. Because

of counterbalancing, however, this effect was orthogonal to our variables of interest, so we do not discuss it further.

Color Word Interest Period.

Target fixations.—We observed an interaction between contrastive accenting and contrast type (see Table 22). To further explore this interaction, we conducted Tukey HSD-corrected post-hoc tests; however, none of them reached significance. Examination of cell means revealed that, during the color word of color-contrast referring expressions, more target fixations occurred when contrastive accenting was present ($M = 0.010$; $SD = 0.100$) than when it was absent ($M = 0.004$; $SD = 0.062$), as is evident in the slightly steeper slope of the purple and red than the blue and black lines during the color word IP of the top left panel of Figure 6. Likewise, during the color word of color- + shape-difference referring expressions, more target fixations occurred when contrastive accenting was present ($M = 0.014$; $SD = 0.118$) than when contrastive accenting was absent ($M = 0.012$; $SD = 0.107$), as is evident in the difference between the purple and red vs. the blue and black lines during the color word IP of the bottom left panel of Figure 6, although this difference was smaller for color-+shape-difference than color-contrast referring expressions. This interaction was qualified by a differently-signed interaction with trial, indicating that these differences decreased over the course of the experiment. These findings suggest that contrastive accenting encouraged anticipation of target referents for both color-contrast and color- + shape-difference critical referring expressions.

Competitor fixations.—We again observed an interaction between contrastive accenting and contrast type (see Table 23). To further explore this interaction, we conducted Tukey HSD-corrected post-hoc tests; however, none of them reached significance. Examination of cell means revealed that, during the color word of color- + shape-difference referring expressions, more fixations of color-contrast competitors occurred when contrastive accenting was present ($M = 0.017$; $SD = 0.128$) than when it was absent ($M = 0.015$; $SD = 0.123$), as is evident in the steeper slope of the purple and red lines than the blue and black lines in the color word IP of the bottom right panel of Figure 6. Moreover, during the color word of color-contrast referring expressions, more fixations of color- + shape-difference competitors occurred when contrastive accenting was present ($M = 0.014$; $SD = 0.118$) than when it was absent ($M = 0.012$; $SD = 0.107$), as is evident in the height of the purple and red lines relative to the blue and black lines in the color word IP of the top right panel of Figure 6. This interaction was also qualified by a differently-signed interaction with trial, indicating that these differences decreased over the course of the experiment. These findings suggest that contrastive accenting encouraged anticipation of competitor referents for both color-contrast and color- + shape-difference critical referring expressions.

Shape Word Interest Period.

Target fixations.—We observed an interaction between beat gesture and contrast type (see Table 24). To further explore this interaction, we conducted Tukey HSD-corrected post-hoc tests; however, none of them reached significance. Examination of cell means revealed that, during the shape word of color- + shape difference critical referring expressions, more target fixations occurred when beat gesture was present ($M = 0.120$; $SD = 0.107$) than when beat

gesture was absent ($M = 0.090$; $SD = 0.095$), as is evident in the steeper slopes of the purple and blue lines relative to the red and black lines in the shape word IP of the bottom left panel of Figure 6. Likewise, during the shape word of color-contrast critical referring expressions, more target fixations occurred when beat gesture was present ($M = 0.100$; $SD = 0.101$) than when beat gesture was absent ($M = 0.080$; $SD = 0.087$), although this difference was slightly smaller, as is evident in the greater overlap between the purple and blue lines and the red and black lines in color word IP of the top left panel of Figure 6. This interaction was also qualified by a differently-signed interaction with trial, indicating that these differences decreased over the course of the experiment. These results indicate that beat gesture accompanying the color word facilitated resolution of critical referring expressions that both contrasted specifically in color and that differed in both color and shape.

Competitor fixations.—We again observed an interaction between beat gesture and contrast type (see Table 26). To further explore this interaction, we conducted Tukey HSD-corrected post-hoc tests; however, none of them reached significance. Examination of cell means revealed that more fixations during the shape word occurred on the color-contrast competitor of color- + shape difference critical referring expressions when beat gesture was present ($M = 0.130$; $SD = 0.155$) than when beat gesture was absent ($M = 0.100$; $SD = 0.101$), as is evident in the higher combined starting points of the purple and blue lines than the red and black lines within the shape word IP of the bottom right panel of Figure 6. Moreover, more fixations during the shape word occurred on the color- + shape difference competitor of color-contrast critical referring expressions when beat gesture was present ($M = 0.080$; $SD = 0.101$) than when beat gesture was absent ($M = 0.070$; $SD = 0.118$), as is evident in the slightly steeper slopes of the purple and blue lines relative to the red and black lines within the shape word IP of the top right panel of Figure 6, although this difference was smaller for color- + shape-difference than color-contrast expressions. This interaction was also qualified by a differently-signed interaction with trial, indicating that these differences decreased over the course of the experiment. These findings indicate that beat gesture accompanying the color word encouraged consideration of competitor referents for both color-contrast and color- + shape-difference critical referring expressions.

Discussion

In Experiment 2, we examined how beat gesture and contrastive accenting affect online processing of critical referring expressions that, relative to a context referring expression, either contrasted in color, such that these cues were locally felicitous, or differed in both color and shape, such that these cues were locally infelicitous. The critical difference from Experiment 1 is that we now tested these variables in a linguistic context in which beat gesture and contrastive accenting were globally infelicitous, given that they never co-occurred with contrastive information in filler referring expressions, and were therefore unreliable cues to contrast across trials. Unlike in Experiment 1, neither beat gesture nor contrastive accenting consistently elicited a contrastive reading. Rather, these cues appeared to encourage consideration of target and competitor referents that both contrasted specifically in color and that differed in both color and shape relative to context referents. These findings support our predictions that global infelicity of beat gesture and contrastive

accenting in Experiment 2 would diminish or eliminate the increases in fixations to the color-contrast competitor of color- + shape-difference critical referring expressions elicited by these cues in Experiment 1, in which these cues were globally felicitous and were reliable cues to contrast across trials.

It is worth noting that the timing of the effects of beat gesture and contrastive accenting on online processing of contrastive information in Experiment 2 differed from those observed in Experiment 1. Specifically, in Experiment 2, contrastive accenting affected anticipatory fixations and beat gesture affected fixations during reference resolution, whereas in Experiment 1, beat gesture affected anticipatory fixations and contrastive accenting affected fixations during reference resolution. These differences in timing suggest that comprehenders may attend to beat gesture before attending to contrastive accenting in a context in which these cues reliably convey contrast, whereas they may attend to contrastive accenting before attending to beat gesture in a context in which these cues fail to reliably convey contrast. This may be the case because comprehenders may rely on their real-world experience interpreting beat gesture, which is typically functionally constrained to conveying contrast, and contrastive accenting, which conveys prominence more generally, to process them in a linguistic context in which they are used atypically. Despite the differences in the timing, however, these cues' facilitation of target and competitor referent identification for both color-contrast and color- + shape-difference critical referring expressions and the decreases in these effects over the course of the experiment indicate that comprehenders attend to how these cues are used within specific linguistic contexts and adjust their interpretation of them accordingly.

Combined Analysis: Experiments 1 and 2

Because the designs of Experiments 1 and 2 were largely similar except for beat gesture and contrastive accent felicity in filler sentences, we conducted combined analyses that included experiment as a fixed effect to directly test the effect of global felicity of beat gesture and contrastive accent with contrast in fillers on reference resolution. In these analyses, target and competitor fixation data was analyzed using the same procedures as in Experiments 1 and 2. All data and analysis scripts are publicly available via the following link: <https://osf.io/qx9tf/>, and numerical results and full outcome descriptions are available in the Supplementary Material section. Here, we report brief descriptions of the results of each analysis, highlighting key main effects and interactions and describing their meanings.

Trial Onset Interest Period

Target fixations.—We observed a significant main effect of experiment (see Table S9), indicating that more fixations on targets occurred during the trial onset interest period of critical referring expressions in Experiment 1 ($M = 0.022$, $SD = 0.151$) than Experiment 2 ($M = 0.001$, $SD = 0.018$). These results indicate that global felicity of beat gesture and contrastive accenting in filler items affected the predisposition to fixate target referents.

We also observed a main effect of contrastive accenting, indicating more fixations on targets during the trial onset interest period occurred when contrastive accenting was absent ($M = 0.012$, $SD = 0.110$) than when it was present ($M = 0.010$, $SD = 0.108$). This effect was

qualified by an interaction with contrast type, which we explored by conducting Tukey HSD-corrected post-hoc tests; however, none of them reached significance. Furthermore, this interaction was qualified by a differently-signed interaction with trial, indicating that it decreased over the course of the experiment. Although it is unclear why these interactions reached significance in this combined analysis in light of evidence that these interactions failed to reach significance in separate analyses by experiment, it is likely that they are spurious; thus, we do not interpret or discuss them further.

Competitor fixations.

We also observed a significant main effect of experiment (see Table S10), indicating that more fixations occurred on competitor objects during the trial onset interest period of critical referring expressions in Experiment 2 ($M = 0.022$, $SD = 0.151$) than Experiment 1 ($M = 0.002$, $SD = 0.044$). These results indicate that global felicity of beat gesture and contrastive accenting in filler items also affected the predisposition to fixate competitor referents.

Color Word Interest Period

Target fixations.—We observed a significant two-way interaction between contrast type and experiment (see Table S11). To further explore this interaction, we conducted Tukey HSD-corrected post-hoc tests; however, none of them reached significance. This interaction was qualified by a differently-signed interaction with trial, indicating that these differences decreased over the course of the experiment.

We also observed a significant interaction between beat gesture and contrast type. This interaction was further qualified by an interaction with experiment, which we explored by conducting Tukey HSD-corrected post-hoc tests (see Table S12). These tests revealed that, during color- + shape-difference critical referring expressions in Experiment 1, more target fixations occurred when beat gesture was absent than when beat gesture was present (see Table S12). That is, within a linguistic context in which beat gesture was globally felicitous with contrast, when critical referring expressions differed in both color and shape from a previous context referring expression, emphasizing the color word with beat gesture reduced anticipatory fixations towards its referent. This interaction was qualified by a differently-signed interaction with trial, indicating that this effect decreased over the course of the experiment.

Competitor fixations.—We observed a significant two-way interaction between contrast type and experiment (see Table S13). Tukey HSD-corrected post-hoc tests revealed that, in Experiment 1, more fixations occurred on color-contrast competitors of color- + shape-difference critical referring expressions than on color- + shape-difference competitors of color-contrast critical referring expressions (see Table S14). This interaction was qualified by a differently-signed interaction with trial, indicating that these differences decreased over the course of the experiment.

We also observed a significant interaction between beat gesture and contrast type. This interaction was further qualified by an interaction with experiment, which we explored by conducting Tukey HSD-corrected post-hoc tests (see Table S15). These tests revealed

that, in Experiment 1, more fixations occurred on color-contrast competitors of color- + shape-difference critical referring expressions when beat gesture was present than when beat gesture was absent (see Table S15). That is, within a linguistic context in which beat gesture was globally felicitous with contrast, the presence of beat gesture during a color word misled comprehenders into anticipating the color-contrast competitor of critical referring expressions differing in both color and shape from a previous context referring expression.

In addition to the effects of the main factors of interest, we also observed a main effect of object side, indicating that, across experiments, more fixations occurred on competitor objects that appeared on one side of the screen than the other.

Shape Word Interest Period

Target fixations.—We observed a significant main effect of experiment (see Table S16), indicating that target fixations occurred more often during the shape word of critical referring expressions of Experiment 1 ($M = 1.010$, $SD = 0.727$) compared to Experiment 2 ($M = 0.010$, $SD = 0.098$). In addition, we observed a significant main effect of contrast type across experiments, indicating that more target fixations occurred during color-contrast critical referring expressions ($M = 0.555$, $SD = 0.756$) than during color- + shape-difference critical referring expressions ($M = 0.478$, $SD = 0.686$). These main effects interacted, which we explored by conducting Tukey HSD-corrected post-hoc tests (see Table S17). These tests revealed that, in Experiment 1, more target fixations occurred during the shape word of color-contrast critical referring expressions than color + shape difference critical referring expressions. Furthermore, they revealed that more target fixations occurred during the shape word of color- + shape-difference critical referring expressions in Experiment 1 than Experiment 2.

We observed a significant three-way interaction of contrastive accenting, beat gesture, and contrast type. Tukey HSD-corrected post-hoc tests revealed that neither beat gesture nor contrastive accenting differentially affected interpretation of color-contrast and color- + shape-difference critical referring expressions (although the reverse was true; see Table S18). In addition, we observed a significant three-way interaction of contrastive accenting, beat gesture, and experiment, which we explored by conducting Tukey HSD-corrected post-hoc tests; however, none of them reached significance. This interaction was qualified by a differently-signed interaction with trial, indicating that these differences decreased over the course of the experiment.

In addition to the effects of the main factors of interest, we observed main effects of gesture orientation and target object side as well as an interaction between them for target fixations during the shape word interest period. This result suggests that more fixations occurred on targets appearing on the side of the array congruent with the orientation of beat gestures when processing shape words in critical referring expressions.

Competitor fixations.

We observed a significant main effect of experiment (see Table S19), indicating that competitor fixations occurred more during the shape word of critical referring expressions

in Experiment 1 ($M = 0.337$, $SD = 0.560$) compared to Experiment 2 ($M = 0.013$, $SD = 0.121$). In addition, we observed a significant main effect of contrastive accenting, indicating that, across experiments, competitor fixations occurred more during the shape word when contrastive accenting was present on the color word of critical referring expressions ($M = 0.187$, $SD = 0.445$) than when it was absent ($M = 0.167$, $SD = 0.432$). Moreover, we observed a significant two-way interaction between contrastive accenting and beat gesture across experiments, which we explored by conducting Tukey HSD-corrected post-hoc tests; however, none of them reached significance.

We also observed a significant two-way interaction between contrastive accenting and contrast. This interaction was qualified by an interaction with experiment, which we explored by conducting Tukey HSD-corrected post-hoc tests (see Table S20). These tests revealed that more fixations of the color-contrast competitor of color + shape difference critical referring expressions occurred in Experiment 1 when contrastive accenting was present on the color word than when contrastive accenting was absent, whereas more fixations on this competitor occurred in Experiment 2 when contrastive accenting was absent than when it was present on the color word. By contrast, more fixations of the color + shape difference competitor of color-contrast critical referring expressions occurred when contrastive accenting was present on the color word than when it was absent in both Experiment 1 and Experiment 2. These results show that, when contrastive accenting is globally felicitous with contrast in filler items, it is interpreted as a cue to contrast when it occurs in critical referring expressions, misleading comprehenders into considering the color-contrast competitor during resolution of critical referring expressions differing in both color and shape from a previous context referring expression.

In addition to the effects of the main factors of interest, we observed main effects of gesture orientation and target object side as well as an interaction between them for fixations on competitor objects during the shape word interest period. This result suggests that more fixations occurred on competitor objects appearing on the side of the array congruent with the orientation of beat gestures when processing shape words in critical referring expressions.

General Discussion

The present research used eye-tracking to examine how the reliability with which two key multimodal cues to prominence—beat gesture and contrastive accenting—indicate contrast affects online processing of contrastive information in spoken discourse. In both Experiments 1 and 2, beat gesture and contrastive accenting were locally felicitous with contrast in half of critical sentences and locally infelicitous with contrast in the other half of critical sentences. In Experiment 1, beat gesture and contrastive accenting always co-occurred with contrast in filler items, such that they were globally felicitous and therefore reliable cues to contrast across trials. In this experiment, when target referents differed in both color and shape from the previous referent, beat gesture or contrastive accenting accompanying the color word of critical referring expressions led to erroneous fixations to competitors specifically contrasting in color with the prior referent. The time course of this effect differed slightly across cues: it first appeared in anticipatory fixations for beat gesture

but only after lexical disambiguation for contrastive accenting. In Experiment 2, in which beat gesture and contrastive accenting never co-occurred with contrast in filler items, such that they were globally infelicitous and were therefore unreliable cues to contrast across trials, beat gesture and contrastive accenting increased fixations on target and competitor referents for both color-contrast and color- + shape-difference critical referring expressions. The time course of this effect again differed across cues, but in a manner that differed from Experiment 1: here, it first appeared in anticipatory fixations for contrastive accenting but only after lexical disambiguation for beat gesture. Taken together, these results indicate that the frequency with which contrastive accenting and beat gesture co-occur with contrastive information influences their interpretation as cues to contrast, in turn affecting online reference resolution in spoken discourse processing.

Effects of Contrastive Accenting

In Experiment 1, in which contrastive accenting reliably indicated contrast in filler items, we found that when the critical referent did not specifically contrast in color with the context referent, contrastive accenting on the color adjective misdirected participants' visual attention to a color-contrast competitor. This result complements prior findings demonstrating that contrastive accenting directs eye movements to contextual contrasts during online reference resolution (Ito & Speer, 2008; Watson et al. 2008). The converse was *not* true: When there was a genuine color contrast, the absence of contrastive accenting on the color adjective did not affect eye movements. This asymmetry has also been obtained elsewhere (Watson et al., 2008; Ito & Speer, 2008) and has been interpreted in conjunction with contrastive accenting's effect on eye movements as indicating that contrastive accenting is more marked than non-contrastive accenting: A contrastive accent obligatorily indicates a contrast and is misleading when not used to emphasize contrastive information, but a non-contrastive accent does not necessarily indicate the absence of a contrast and so does not hinder comprehension even when it occurs in a contrastive context.

Unlike this previous work, however, although the presence of infelicitous contrastive accenting without a genuine color contrast misdirected eye movements, the *felicitous* presence of contrastive accenting on a contrasting color adjective failed to facilitate online reference resolution. One possible explanation for this difference in results is the presence of a salient video stimulus in the current work but not this previous work. More specifically, in the current work, beat gesture, which also reliably indicated contrast in filler items in Experiment 1, was varied orthogonally with contrastive accenting on color adjectives of critical items. Thus, it may have been interpreted as more salient (yet equally reliable) relative to contrastive accenting within this context, as suggested by other previous research in which both cues occur and are manipulated relative to one another (Kushch et al., 2018; Kushch & Prieto, 2016; Llanes-Coromina et al., 2018; Morett & Fraundorf, 2019). To investigate this possibility, future research should examine the impact of contrastive accenting on online contrast interpretation in linguistic contexts in which both contrastive accenting and beat gesture occur vs. only contrastive accenting occurs when these cues indicate contrast with equal reliability.

With respect to effect timing, in Experiment 1, contrastive accenting reliably affected competitor fixations during the shape word, indicating that it influenced reference resolution after disambiguating lexical information had arrived, but—unlike beat gesture—*not* anticipation of referents in advance of disambiguating lexical input. By contrast, in Experiment 2, in which contrastive accenting failed to reliably indicate contrast in filler items, contrastive accenting affected anticipation of target and competitor referents in advance of disambiguating lexical input, and this effect decreased in magnitude over the course of the experiment. Prior visual world experiments with healthy adults have yielded mixed results about the timing of contrastive accenting effects on spoken language processing: Some have revealed anticipatory effects (Watson et al., 2008), but others have not (Ito & Speer, 2008). The contrastive effects that we observed in Experiment 1 align with other experiments that did not find anticipatory contrastive effects, which may reflect similarity of stimuli: As in Ito and Speer (2008), our critical sentences consisted of a command referring to an object differing in shape and/or color from the object mentioned in the preceding context, as opposed to a command referring to the location of an object previously mentioned in preceding discourse (Watson et al., 2008).

Effects of Beat Gesture

The finding that beat gesture differentially influences fixations during interpretation of color-contrast and color- + shape-difference critical referring expressions in Experiments 1 and 2 constitutes the first evidence that beat gesture affects online reference resolution during spoken language processing. More specifically, the evidence for a contrastive interpretation of beat gesture in Experiment 1, in which beat gesture was globally felicitous and was therefore a reliable cue to contrast across trials, complements evidence that beat gesture accompanying selected items in a contrastive set introduced in discourse enhances subsequent memory for these items (Kushch et al., 2018; Llanes-Coromina et al., 2018; Morett & Fraundorf, 2019). In Experiment 2, in which beat gesture was globally infelicitous and was therefore an unreliable cue to contrast across trials, the absence of a similar contrastive interpretation of beat gesture is consistent with research demonstrating that interpretation of cues to contrast—including beat gesture—is influenced by the reliability with which they indicate contrast elsewhere within a global linguistic context (see also Grodner & Sedivy, 2011; Morett & Fraundorf, 2019; Roettger & Franke, 2019; Roettger & Rimland, 2020; Ryskin et al., 2019). Although these findings suggest that beat gesture and speech mutually influence one another, as postulated by the integrated systems hypothesis (Kelly et al., 2010), future research should investigate the extent to which this integration is obligatory (another key aspect of this hypothesis) by examining whether beat gesture and speech influence one another when comprehenders are instructed to ignore cues in each modality when both cues are present.

As the first research to employ the visual world eye-tracking paradigm to examine beat gesture-speech integration, the results provide important insight into the timing of beat gesture's influence on online spoken language comprehension. In Experiment 1, beat gesture first affected fixations during the color word period and persisted into the shape word period of critical referring expressions, indicating that it influenced reference anticipation as well as reference resolution in this context. By contrast, in Experiment 2, beat gesture only affected

fixations during the shape word period, indicating that its effects were constrained to reference resolution and did not affect reference anticipation in this context. These findings are consistent with previous research demonstrating an immediate and persistent influence of iconic gesture on online spoken language interpretation in healthy adolescents within a supportive linguistic context (Silverman et al., 2010), demonstrating that gesture-speech integration occurs instantaneously and persists through the lexical affiliate in such a context. In the current research, the apex of beat gesture preceded stressed syllables of color words by 200 ms, with preparatory hand movement preceding the color word, which in principle would have allowed comprehenders to interpret beat gesture as an anticipatory cue to the upcoming referent. To more clearly illuminate similarities and differences between beat and iconic gesture in their timing and influence on online spoken language interpretation, future research should directly compare the effects of these types of gestures using a visual world paradigm in which they accompany similar discourse and are temporally aligned with respect to their strokes.

Comparing Beat Gesture and Contrastive Accenting

Broadly, our results suggest that beat gesture and contrastive accenting exert independent influences on online spoken discourse processing insofar as (a) the effects of beat gesture and contrastive accenting had different time courses and (b) did not interact. Notably, these differences between beat gesture and contrastive accenting were present in both experiments of the current research; the varying validity of these cues as indicators of contrast across experiments did not affect (the absence of) online integration of these cues, despite that the time courses of their effects differed between experiments. The absence of interactive effects of beat gesture and contrastive accenting observed in the current work differs from previous research demonstrating such effects on memory for spoken language, in which the two cues were observed to interact (Kushch et al., 2018; Kushch & Prieto, 2016; Llanes-Coromina et al., 2018; Morett & Fraundorf, 2019), possibly reflecting a difference between online and offline measures of cue interpretation. Moreover, it suggests that beat gesture and contrastive accenting may be integrated over an extended time period rather than instantaneously.

Cue Validity in a Linguistic Context

Comparing results across our two visual world eye-tracking experiments demonstrates that online interpretation of cues to reference resolution can be modulated by how those cues are used within the present linguistic context (Grodner & Sedivy, 2011; Kleinschmidt et al., 2012; Roettger & Franke, 2019; Roettger & Rimland, 2020; Ryskin et al., 2019). Specifically, contrastive accenting encouraged garden-path fixations to color-contrast competitor referents of color- + shape-difference critical referring expressions in Experiment 1 but not Experiment 2, suggesting that contrastive accenting is more likely to be interpreted as a cue to contrast when it generally occurs felicitously (Experiment 1) rather than infelicitously (Experiment 2) with contrast. These findings are consistent with previous findings indicating that listeners alter their interpretation of contrastive accenting based on whether or not speakers use it felicitously as a cue to contrast (Roettger & Franke, 2019; Roettger & Rimland, 2020). We also observed differences across experiments in interpretation of beat gesture: Beat gesture elicited a similar garden-path effect in Experiment 1 but not Experiment 2, in which it encouraged consideration of competitor

referents for both color-contrast and color- + shape-difference critical referring expressions. Thus, the general pattern of findings indicates that both beat gesture and contrastive accenting were interpreted as cues to contrast during processing of critical referring expressions in Experiment 1 but not Experiment 2. These findings demonstrate that the overall reliability with which beat gesture and pitch accenting convey contrast affects how these cues are interpreted during online spoken discourse processing.

Future Directions

Although the results of the current research provide an important first glimpse into how beat gesture and contrastive accenting influence online spoken discourse interpretation, additional research is needed to provide a more comprehensive view of this issue. In particular, to more clearly distinguish between beat gesture's influence on contrast interpretation versus its influence on salience more generally, it will be necessary to conduct visual world eye-tracking research in which beat gesture is manipulated with respect to both contrast and salience in spoken discourse. Moreover, although our results provide evidence that comprehenders are sensitive to the contingencies between linguistic cues and referential intent within specific linguistic contexts, this claim could be further tested by varying the reliability and validity of other linguistic cues (e.g., focus, pronouns) to assess how this affects their online interpretation. Finally, to demonstrate robustness across methods, similar questions could be investigated using alternate measure of online spoken language processing, such as mouse tracking.

Conclusion

The results of the current research demonstrate that beat gesture and contrastive accenting independently influence contrast interpretation during online reference resolution in spoken discourse comprehension. Both of these cues can be interpreted as cues to contrast. However, online interpretation of these cues is guided by whether they reliably or unreliably convey contrast within the current discourse context. Taken together, these findings indicate that the felicity of beat gesture and contrastive accenting with contrastive information within different discursive contexts affects whether each of these cues is interpreted contrastively.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by a Hilibrand Postdoctoral Fellowship to L.M.M. and by award MH107426 from the National Institutes of Mental Health to J.C.M. All experimental stimuli, analysis scripts, and data are publicly available via the Open Science Framework at <https://osf.io/qx9tf/>. The authors thank Simone Hasselmo and Julie Trapani for assistance with auditory and video recording; Ellen Macaruso, Jake Feiler, Sarah Hughes Berheim, Cailee Nelson, Nan Mu, and Nathaniel Shannon IV for assistance with stimulus preparation and paradigm testing; Talena Day, Kathryn McNaughton, Autumn Christafore, Julia Hahn, and Alyssa Schoonmaker for assistance with data collection; and Maggie Paul for assistance with the norming study. In addition, the authors thank the attendees of the 2018 meetings of the CUNY Conference on Human Sentence Processing and the Psychonomic Society and the 2019 meeting of the Cognitive Science Society for their feedback on portions of this research.

References

- Audacity Team. (2013). Audacity (2.0.3) [Computer software]. <http://audacityteam.org/>
- Austin EE, & Sweller N. (2014). Presentation and production: The role of gesture in spatial communication. *Journal of Experimental Child Psychology*, 122, 92–103. 10.1016/j.jecp.2013.12.008 [PubMed: 24549229]
- Barr DJ (2008). Analyzing ‘visual world’ eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, 59, 457–474.
- Bates D, Mächler M, Bolker B, & Walker S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48. 10.18637/jss.v067.i01
- Biau E, & Soto-Faraco S. (2013). Beat gestures modulate auditory integration in speech perception. *Brain and Language*, 124(2), 143–152. 10.1016/j.bandl.2012.10.008 [PubMed: 23333667]
- Biau E, Torralba M, Fuentemilla L, de Diego Balaguer R, & Soto-Faraco S. (2015). Speaker’s hand gestures modulate speech perception through phase resetting of ongoing neural oscillations. *Cortex*, 68, 76–85. 10.1016/j.cortex.2014.11.018 [PubMed: 25595613]
- Boersma P, & Weenink D. (2016). Praat: Doing phonetics by computer (6.0.39) [Computer software].
- Braun B, & Tagliapietra L. (2010). The role of contrastive intonation contours in the retrieval of contextual alternatives. *Language and Cognitive Processes*, 25, 1024–1043. 10.1080/01690960903036836
- Christianson K, Hollingworth A, Halliwell JF, & Ferreira F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology*, 42(4), 368–407. [PubMed: 11368528]
- Dimitrova D, Chu M, Wang L, Özyürek A, & Hagoort P. (2016). Beat that word: How listeners integrate beat gesture and focus in multimodal speech discourse. *Journal of Cognitive Neuroscience*, 28(9), 1255–1269. 10.1162/jocn_a_00963 [PubMed: 27027421]
- Esteve-Gibert N, & Prieto P. (2013). Prosodic structure shapes the temporal realization of intonation and manual gesture movements. *Journal of Speech, Language, and Hearing Research*, 56(3), 850–864. 10.1044/1092-4388(2012/12-0049)
- Feyereisen P. (2006). Further investigation on the mnemonic effect of gestures: Their meaning matters. *European Journal of Cognitive Psychology*, 18, 185–205. 10.1080/09541440540000158
- Fraundorf SH, Benjamin AS, & Watson DG (2013). What happened (and what did not): Discourse constraints on encoding of plausible alternatives. *Journal of Memory and Language*, 69, 196–227. 10.1016/j.jml.2013.06.003 [PubMed: 24014934]
- Fraundorf SH, Watson DG, & Benjamin AS (2010). Recognition memory reveals just how CONTRASTIVE contrastive accenting really is. *Journal of Memory and Language*, 63, 367–386. 10.1016/j.jml.2010.06.004 [PubMed: 20835405]
- Fraundorf SH, Watson DG, & Benjamin AS (2012). The effects of age on the strategic use of pitch accents in memory for discourse: A processing-resource account. *Psychology and Aging*, 27, 88–98. 10.1037/a0024138 [PubMed: 21639646]
- Gluhareva D, & Prieto P. (2017). Training with rhythmic beat gestures benefits L2 pronunciation in discourse-demanding situations. *Language Teaching Research*, 21, 609–631. 10.1177/1362168816651463
- Gotzner N, & Spalek K. (2019). The life and times of focus alternatives: Tracing the activation of alternatives to a focused constituent in language comprehension. *Language and Linguistics Compass*, 13, e12310. 10.1111/lnc3.12310
- Grodner D, & Sedivy JC (2011). The effect of speaker-specific information on pragmatic inferences. In Pearlmutter N. & Gibson E. (Eds.), *The processing and acquisition of reference* (pp. 239–272). MIT Press.
- Hirata Y, Kelly SD, Huang J, & Manansala M. (2014). Effects of hand gestures on auditory learning of second-language vowel length contrasts. *Journal of Speech, Language, and Hearing Research*, 57, 2090–2101. 10.1044/2014_JSLHR-S-14-0049
- Huang YT, & Arnold AR (2016). Word learning in linguistic context: Processing and memory effects. *Cognition*, 156, 71–87. [PubMed: 27513871]

- Huang YT, & Snedeker J. (2018). Some inferences still take time: Prosody, predictability, and the speed of scalar implicatures. *Cognitive Psychology*, 102, 105–126. [PubMed: 29454819]
- Huetig F. (2015). Four central questions about prediction in language processing. *Brain Research*, 1626, 118–135. 10.1016/j.brainres.2015.02.014 [PubMed: 25708148]
- Husband EM, & Ferreira F. (2016). The role of selection in the comprehension of focus alternatives. *Language, Cognition and Neuroscience*, 31, 217–235. 10.1080/23273798.2015.1083113
- Igualada A, Esteve-Gibert N, & Prieto P. (2017). Beat gestures improve word recall in 3-to 5-year-old children. *Journal of Experimental Child Psychology*, 156, 99–112. 10.1016/j.jecp.2016.11.017 [PubMed: 28061372]
- Ito K, Jincho N, Minai U, Yamane N, & Mazuka R. (2012). Intonation facilitates contrast resolution: Evidence from Japanese adults and 6-year olds. *Journal of Memory and Language*, 66, 265–284. 10.1016/j.jml.2011.09.002
- Ito K, & Speer SR (2008). Anticipatory effects of intonation: Eye movements during instructed visual search. *Journal of Memory and Language*, 58(2), 541–573. 10.1016/j.jml.2007.06.013 [PubMed: 19190719]
- Kelly SD, Özyürek A, & Maris E. (2010). Two sides of the same coin speech and gesture mutually interact to enhance comprehension. *Psychological Science*, 21(2), 260–267. 10.1177/0956797609357327 [PubMed: 20424055]
- Kleinschmidt DF, Fine AB, & Jaeger TF (2012). A belief-updating model of adaptation and cue combination in syntactic comprehension. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 34.
- Krahmer E, & Swerts M. (2007). The effects of visual beats on prosodic prominence: Acoustic analyses, auditory perception and visual perception. *Journal of Memory and Language*, 57(3), 396–414. 10.1016/j.jml.2007.06.005
- Kurumada C, Brown M, Bibyk S, Pontillo DF, & Tanenhaus MK (2014). Is it or isn't it: Listeners make rapid use of prosody to infer speaker meanings. *Cognition*, 133(2), 335–342. 10.1016/j.cognition.2014.05.017 [PubMed: 25128792]
- Kushch O, Igualada A, & Prieto P. (2018). Prominence in speech and gesture favour second language novel word learning. *Language, Cognition and Neuroscience*, 33(8), 992–1004. 10.1080/23273798.2018.1435894
- Kushch O, & Prieto P. (2016). The effects of pitch accentuation and beat gestures on information recall in contrastive discourse. In Barnes J, Brugos A, Shattuck-Hufnagel S, & Veilleux N. (Eds.), *Proceedings of the 8th International Conference on Speech Prosody* (pp. 922–925). International Speech Communication Association.
- Kuznetsova A, Brockhoff PB, & Christensen RHB (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82, 1–26.
- Ladd DR (1996). *Intonational phonology*. Cambridge University Press.
- Lee E-K, & Snedeker J. (2016). Effects of contrastive accents on children's discourse comprehension. *Psychonomic Bulletin & Review*, 23, 1589–1595. 10.3758/s13423-016-1069-7 [PubMed: 27234261]
- Lenth R. (2019). emmeans: Estimated Marginal Means, aka Least-Squares Means (1.4.3) [Computer software]. <https://CRAN.R-project.org/package=emmeans>
- Leonard T, & Cummins F. (2011). The temporal relation between beat gestures and speech. *Language and Cognitive Processes*, 26(10), 1457–1471. 10.1080/01690965.2010.500218
- Levantinou EI, & Navarretta C. (2016). An investigation of the effect of beat and iconic gestures on memory recall in L2 speakers. In *Proceedings from the 3rd European Symposium on Multimodal Communication* (pp. 32–37). Linköping University Electronic Press.
- Llanes-Coromina J, Vilà-Giménez I, Kushch O, Borràs-Comes J, & Prieto P. (2018). Beat gestures help preschoolers recall and comprehend discourse information. *Journal of Experimental Child Psychology*, 172, 168–188. 10.1016/j.jecp.2018.02.004 [PubMed: 29626756]
- Macoun A, & Sweller N. (2016). Listening and watching: The effects of observing gesture on preschoolers' narrative comprehension. *Cognitive Development*, 40, 68–81. 10.1016/j.cogdev.2016.08.005
- McNeill D. (1992). *Hand and mind*. University of Chicago Press.

- McNeill D. (2005). *Gesture and thought*. University of Chicago Press.
- McNeill D. (2006). Gesture: A psycholinguistic approach. *The Encyclopedia of Language and Linguistics*, 58–66.
- Morett LM (2014). When hands speak louder than words: The role of gesture in the communication, encoding, and recall of words in a novel second language. *The Modern Language Journal*, 98, 834–853. 10.1111/modl.12125
- Morett LM, & Fraundorf SH (2020, October 5). Eye see what you're saying: Contrastive use of beat gesture and pitch accent affects online interpretation of spoken discourse. Retrieved from osf.io/qx9tf. 10.17605/OSF.IO/QX9TF
- Morett LM, & Fraundorf SH (2019). Listeners consider alternative speaker productions in discourse comprehension and memory: Evidence from beat gesture and pitch accenting. *Memory & Cognition*, 47(8), 1515–1530. 10.3758/s13421-019-00945-1 [PubMed: 31215010]
- Myhill J, & Xing JZ (1996). Towards an operational definition of discourse contrast. *Studies in Language*, 20(2), 303–360. 10.1075/sl.20.2.04myh
- Pierrehumbert J, & Hirschberg J. (1990). The meaning of intonational contours in the interpretation of discourse. In Cohen PR (Ed.), *Intentions in communication* (pp. 271–311). MIT Press.
- Roettger TB, & Franke M. (2019). Evidential strength of intonational cues and rational adaptation to (un-) reliable intonation. *Cognitive Science*, 43(7), e12745. [PubMed: 31310022]
- Roettger TB, & Rimland K. (2020). Listeners' adaptation to unreliable intonation is speaker-sensitive. *Cognition*, 204, 104372. [PubMed: 32615469]
- Roustan B, & Dohen M. (2010a). Co-production of contrastive prosodic focus and manual gestures: Temporal coordination and effects on the acoustic and articulatory correlates of focus. In Hasegawa-Johnson M. (Ed.), *Speech Prosody 2010*. International Speech Communication Association.
- Roustan B, & Dohen M. (2010b). Gesture and speech coordination: The influence of the relationship between manual gesture and speech. In Hasegawa-Johnson M. (Ed.), *Speech Prosody 2010*. International Speech Communication Association.
- Rusiewicz HL, Shaiman S, Iverson JM, & Szuminsky N. (2013). Effects of prosody and position on the timing of deictic gestures. *Journal of Speech, Language, and Hearing Research*, 56(2), 458–470. 10.1044/1092-4388(2012/11-0283)
- Rusiewicz HL, Shaiman S, Iverson JM, & Szuminsky N. (2014). Effects of perturbation and prosody on the coordination of speech and gesture. *Speech Communication*, 57, 283–300. 10.1016/j.specom.2013.06.004
- Ryskin R, Kurumada C, & Brown-Schmidt S. (2019). Information integration in modulation of pragmatic inferences during online language comprehension. *Cognitive Science*, 43(8), e12769. 10.1111/cogs.12769
- Sanford AJS, Sanford AJ, Molle J, & Emmott C. (2006). Shallow processing and attention capture in written and spoken discourse. *Discourse Processes*, 42, 109–130. 10.1207/s15326950dp4202_2
- Selkirk E. (1995). Sentence prosody: Intonation, stress, and phrasing. In Goldsmith JA (Ed.), *The handbook of phonological theory* (Vol. 1, pp. 550–569). Blackwell.
- Shattuck-Hufnagel S, & Ren A. (2018). The prosodic characteristics of non-referential cospeech gestures in a sample of academic-lecture-style speech. *Frontiers in Psychology*, 9. 10.3389/fpsyg.2018.01514
- Silverman LB, Bennetto L, Campana E, & Tanenhaus MK (2010). Speech-and-gesture integration in high functioning autism. *Cognition*, 115(3), 380–393. 10.1016/j.cognition.2010.01.002 [PubMed: 20356575]
- Slattery TJ, Sturt P, Christianson K, Yoshida M, & Ferreira F. (2013). Lingering misinterpretations of garden path sentences arise from competing syntactic representations. *Journal of Memory and Language*, 69(2), 104–120.
- So WC, Sim Chen-Hui C, & Low Wei-Shan J. (2012). Mnemonic effect of iconic gesture and beat gesture in adults and children: Is meaning in gesture important for memory recall? *Language and Cognitive Processes*, 27(5), 665–681. 10.1080/01690965.2011.573220

- Vilà-Giménez I, Igualada A, & Prieto P. (2019). Observing storytellers who use rhythmic beat gestures improves children's narrative discourse performance. *Developmental Psychology*, 55(2), 250–262. 10.1037/dev0000604 [PubMed: 30489138]
- Wang L, & Chu M. (2013). The role of beat gesture and pitch accent in semantic processing: An ERP study. *Neuropsychologia*, 51(13), 2847–2855. 10.1016/j.neuropsychologia.2013.09.027 [PubMed: 24060845]
- Watson DG, Tanenhaus MK, & Gunlogson CA (2008). Interpreting pitch accents in online comprehension: H* vs. L+ H*. *Cognitive Science*, 32, 1232–1244. 10.1080/03640210802138755 [PubMed: 21585451]
- Weber A, Braun B, & Crocker MW (2006). Finding referents in time: Eye-tracking evidence for the role of contrastive accents. *Language and Speech*, 49, 367–392. 10.1177/00238309060490030301 [PubMed: 17225671]
- Yap D-F, So W-C, Yap MJ-M, Tan Y-Q, & Teoh R-LS (2011). Iconic gestures prime words. *Cognitive Science*, 35, 171–183. 10.1111/j.15516709.2010.01141.x [PubMed: 21428996]

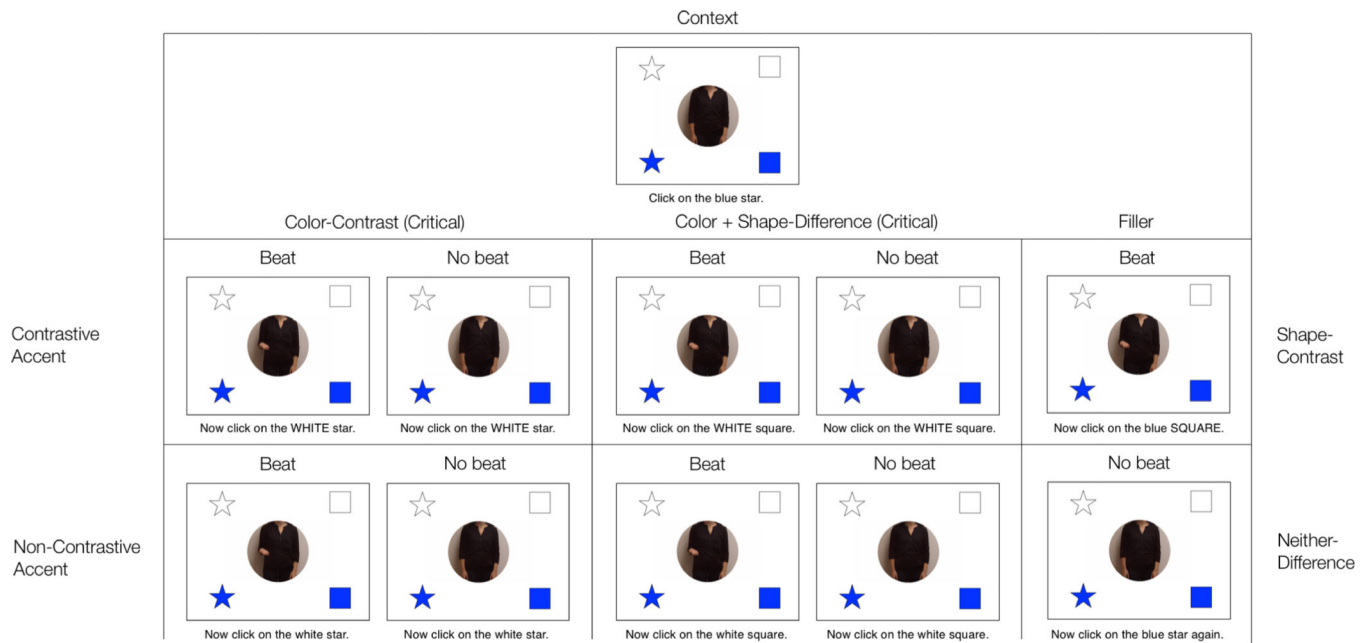


Figure 1. Visual Schematic of Sentences and Accompanying Visual Displays from First Critical and Filler Trials of Lists 1.1–1.4 in Expt. 1

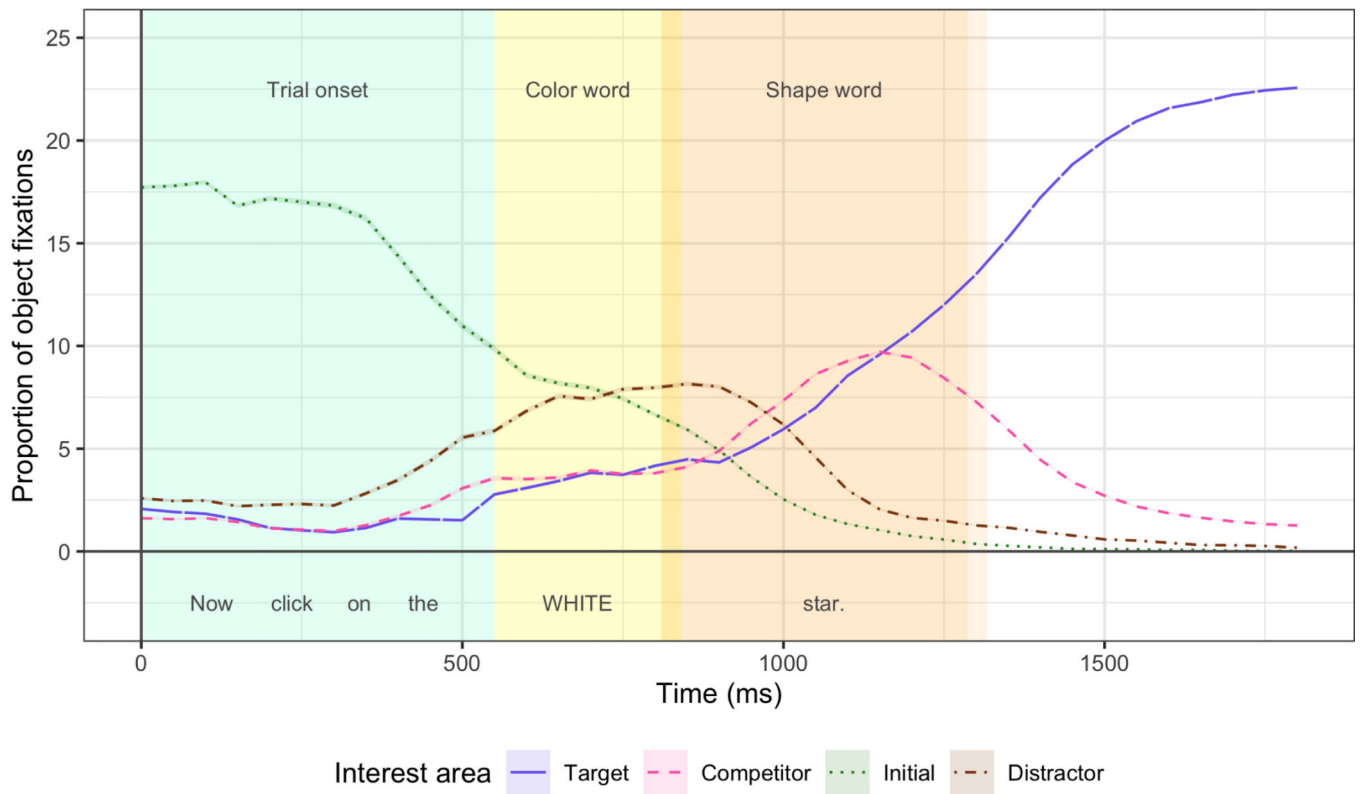


Figure 2. Proportion of Target, Competitor, Initial, and Distractor Fixations Over Time During Critical Referring Expressions in Expt. 1 (Line Shading Represents Standard Error)

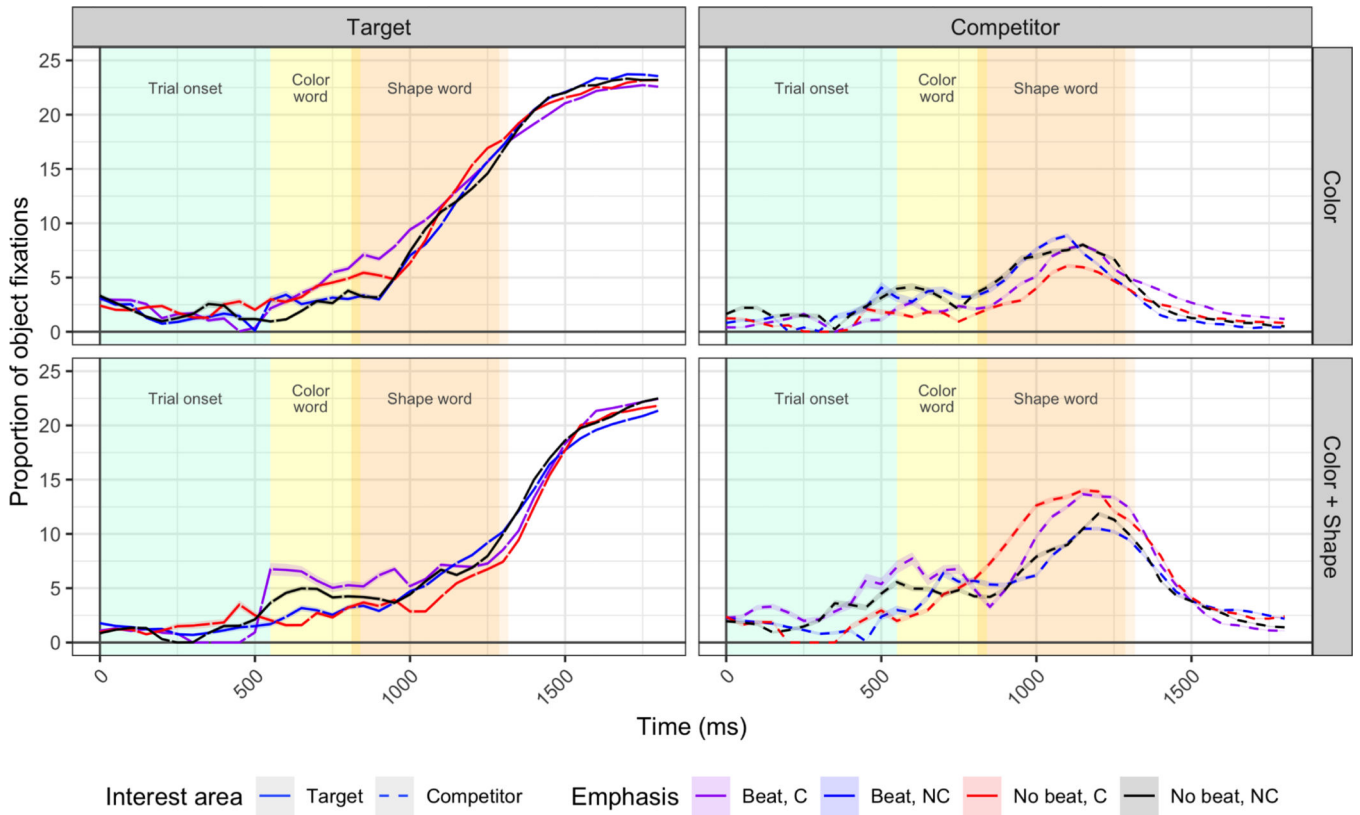


Figure 3. Proportion of Target and Competitor Fixations by Beat Gesture and Pitch Accenting Over Time During Color-Contrast and Color- + Shape-Difference Critical Referring Expressions in Expt. 1 (Line Shading Represents Standard Error; C: Contrastive, NC: Non-Contrastive)

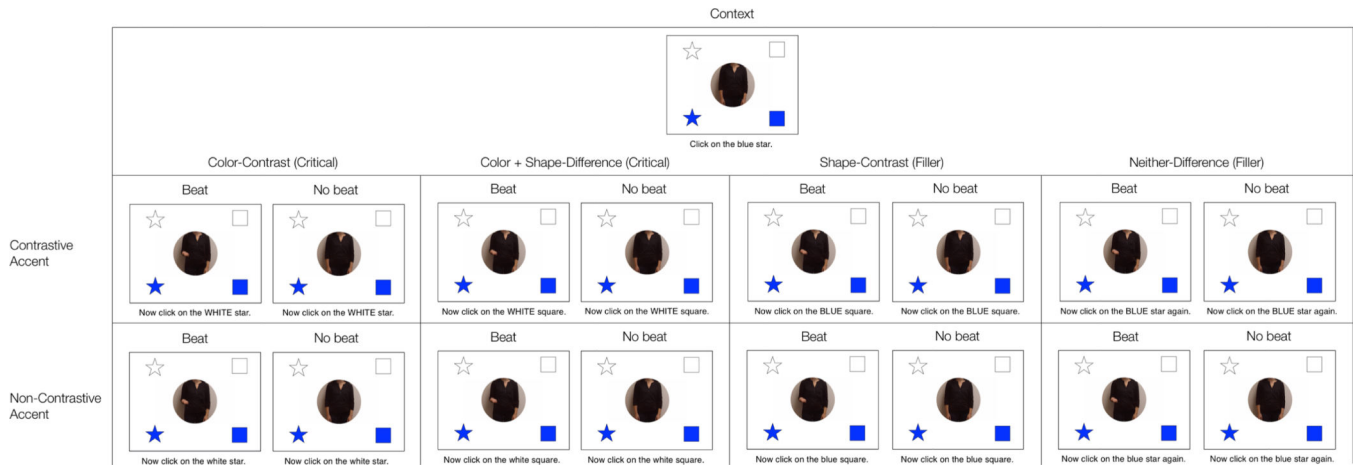


Figure 4. Visual Schematic of Sentences and Accompanying Visual Displays from First Critical and Filler Trials of Lists 1.1–1.4 in Expt. 2

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

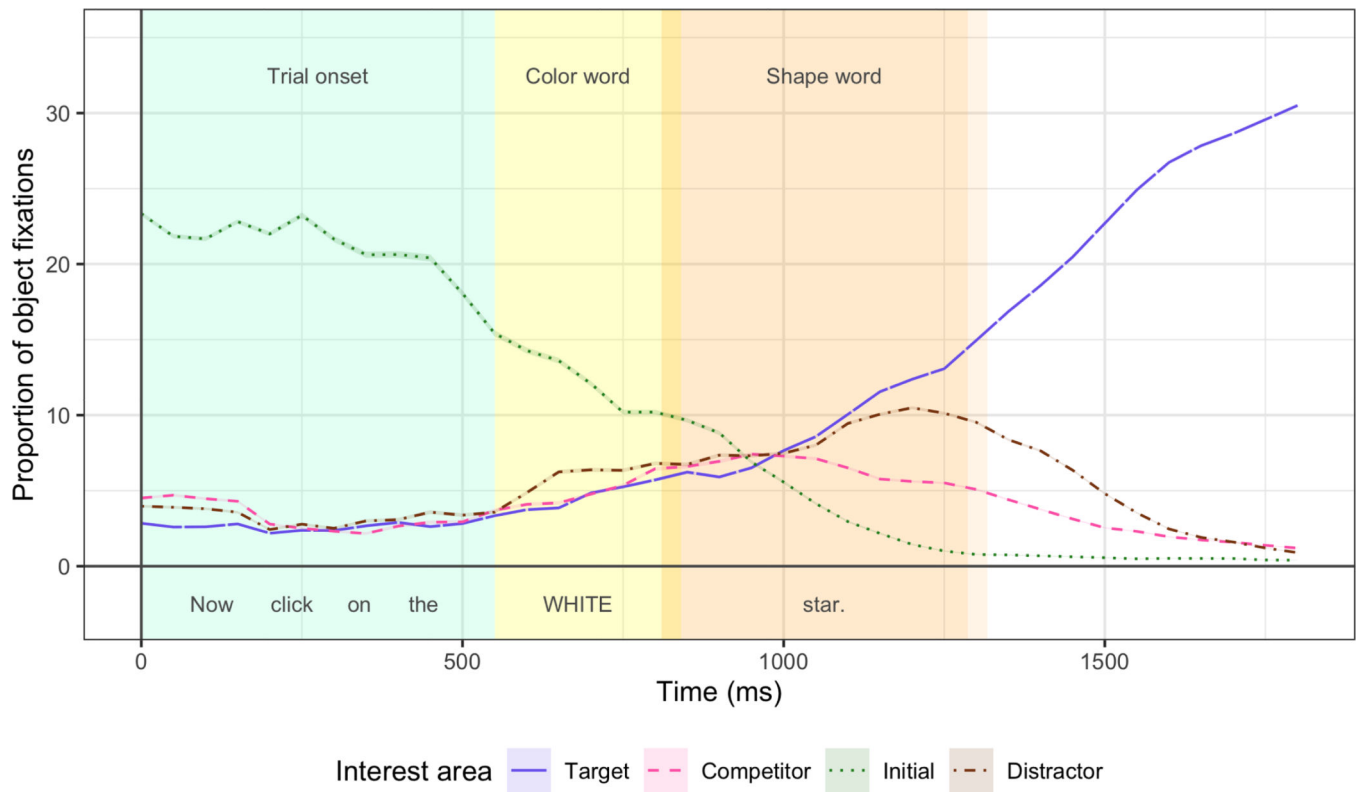


Figure 5. Proportion of Target, Competitor, Initial, and Distractor Fixations Over Time During Critical Referring Expressions in Expt. 2 (Line Shading Represents Standard Error)

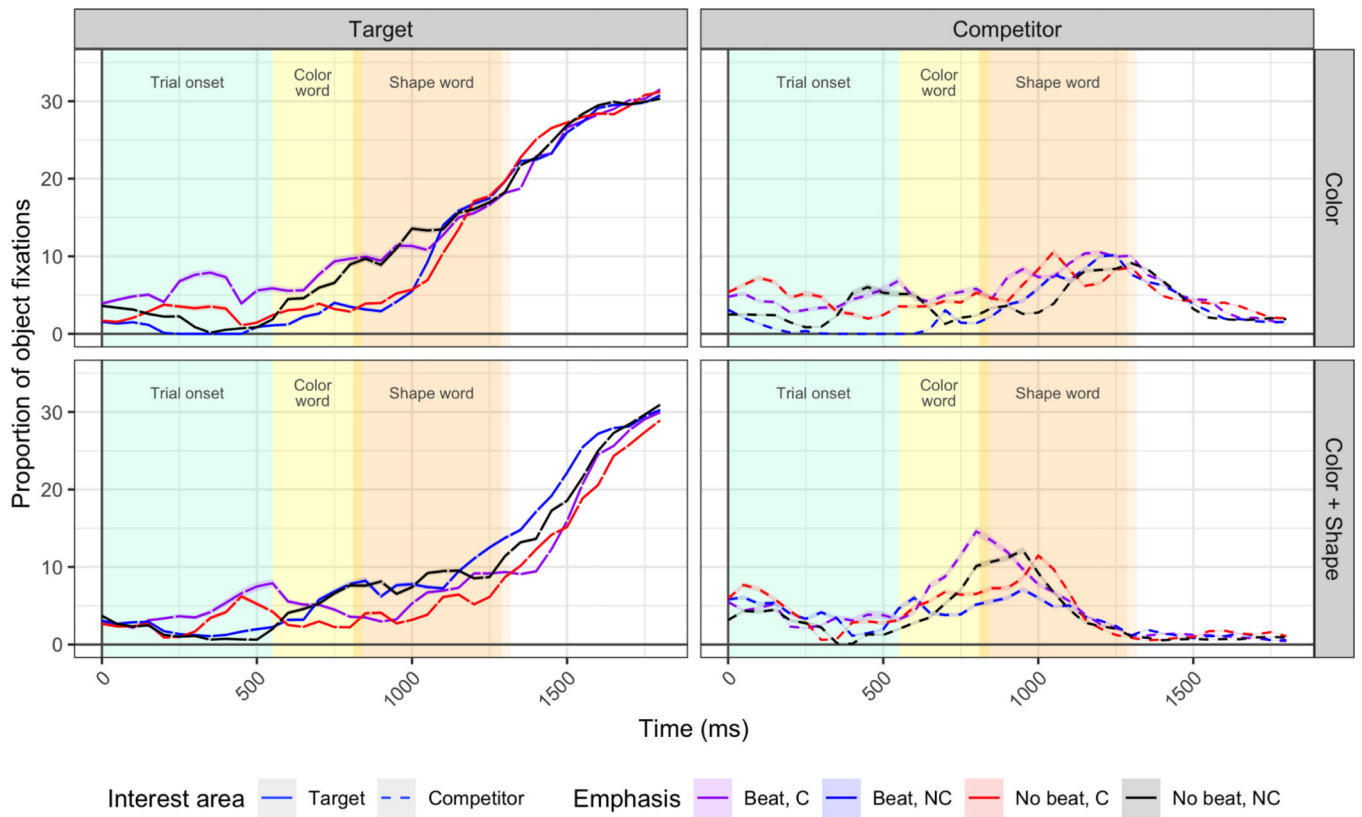


Figure 6. Proportion of Target and Competitor Fixations by Beat Gesture and Pitch Accenting Over Time During Color-Contrast and Color- + Shape-Difference Critical Referring Expressions in Expt. 2 (Line Shading Represents Standard Error; C: Contrastive, NC: Non-Contrastive)

Table 1

Colors and Shapes of Objects Used in Arrays

Colors	Shapes
Red	Circle
Blue	Square
Green	Triangle
Yellow	Rectangle
Orange	Diamond
Purple	Oval
White	Star
Black	Heart

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Acoustic Measures by Accent Type for Color and Shape Words Used in Continuation Sentences

Measure	Non-Contrastive	Contrastive	df	t	p
Critical sentence (color)					
<i>Stressed syllable</i>					
Intensity (dB)	35.33 (1.50)	38.30 (1.66)	58	-7.27	<.001***
Duration (s)	0.15 (0.02)	0.18 (0.03)	58	-4.53	<.001***
Maximum F_0 (Hz)	224.67 (26.17)	330.34 (53.79)	58	-9.68	<.001***
F_0 difference (Hz)	28.41 (17.58)	96.46 (57.35)	58	-6.21	<.001***
Mean F_0 (Hz)	208.53 (21.20)	275.08 (37.26)	58	-8.50	<.001***
<i>Entire word</i>					
Intensity (dB)	35.60 (1.82)	38.04 (1.90)	158	-8.31	<.001***
Duration (s)	0.26 (0.05)	0.29 (0.06)	158	-4.08	<.001***
Maximum F_0 (Hz)	315.45 (141.38)	357.81 (73.33)	158	-2.38	.02*
F_0 difference (Hz)	122.35 (141.30)	146.52 (75.74)	158	-1.35	.18
Mean F_0 (Hz)	217.06 (21.61)	280.89 (21.34)	158	-18.80	<.001***
Filler sentence (shape)					
<i>Stressed syllable</i>					
Intensity (dB)	37.37 (2.00)	39.16 (1.85)	58	-3.62	<.001***
Duration (s)	0.20 (0.02)	0.24 (0.05)	58	-4.83	<.001***
Maximum F_0 (Hz)	370.15 (170.52)	361.29 (117.67)	58	0.23	.82
F_0 difference (Hz)	177.81 (176.45)	124.05 (121.38)	58	1.37	.18
Mean F_0 (Hz)	235.48 (40.05)	278.40 (31.88)	58	-4.59	<.001***
<i>Entire word</i>					
Intensity (dB)	37.24 (1.98)	36.97 (1.81)	158	0.90	.37
Duration (s)	0.34 (0.07)	0.52 (0.09)	158	-13.90	<.001***
Maximum F_0 (Hz)	383.11 (171.64)	435.82 (137.17)	158	-2.15	.03*
F_0 difference (Hz)	208.57 (172.45)	266.63 (147.95)	158	-2.29	.02*
Mean F_0 (Hz)	226.09 (47.56)	261.82 (39.81)	158	-5.15	<.001***

Note:

*
 $p < .05$;***
 $p < .001$

Table 3

Proportion of Fixations to Each Interest Area by Interest Period in Expt. 1

Interest period	Video			Initial	Target	Competitor	Distractor
	Beat	No beat	Overall				
Trial onset	0.87	0.86	0.86	0.09	0.01	0.01	0.02
Color word	0.89	0.85	0.87	0.04	0.03	0.03	0.04
Shape word	0.45	0.46	0.45	0.02	0.36	0.12	0.05

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4

Fixed Effect (Top) and Variance Estimates (Bottom) for Multi-Level Model of Video Fixations During Trial Onset Interest Period in Expt. 1 (Observations = 3,206)

Fixed effect	Coefficient	SE	<i>t</i>	<i>p</i>
Intercept	0.51	0.09	5.94	< .001 ***
Contrastive accenting	-0.10	0.16	-0.62	.54
Beat gesture	0.04	0.04	0.90	.37
Contrast type	0.02	0.03	0.64	.52
Contrastive accenting x beat gesture	0.10	0.08	1.17	.24
Contrastive accenting x contrast type	0.37	0.32	1.15	.26
Beat gesture x contrast type	0.07	0.06	1.12	.26
Contrastive accenting x beat gesture x contrast type	-0.28	0.12	-2.34	.02*
Gesture orientation	-0.05	0.03	-1.72	.09 [†]
Object side	-0.08	0.03	-2.72	.007**
Gesture orientation x object side	0.12	0.06	1.96	.05 [†]

Random effect	<i>s</i> ²
Trial	0.24
Participant	0.49

Note:

* $p < .05$;

** $p < .01$;

*** $p < .001$;

[†] $p < .10$

Table 5

Fixed Effect (Top) and Variance Estimates (Bottom) for Multi-Level Model of Video Fixations During Color Word Interest Period in Expt. 1 (Observations = 3,212)

Fixed effect	Coefficient	SE	t	p
Intercept	-0.88	0.06	-15.29	< .001***
Contrastive accenting	-0.08	0.10	-0.82	.41
Beat gesture	-0.04	0.04	-0.97	.33
Contrast type	0.05	0.03	1.62	.11
Contrastive accenting x beat gesture	0.01	0.08	0.08	.94
Contrastive accenting x contrast type	-0.06	0.19	-0.34	.73
Beat gesture x contrast type	-0.04	0.06	-0.65	.52
Contrastive accenting x beat gesture x contrast type	-0.27	0.12	-2.32	.02*
Gesture orientation	0.10	0.03	3.25	.001**
Object side	0.04	0.03	1.37	.17
Gesture orientation x object side	0.03	0.06	0.42	.67

Random effect	s^2
Trial	0.27
Participant	0.28

Note:

* $p < .05$;

** $p < .01$;

*** $p < .001$

Table 6

Fixed Effect (Top) and Variance Estimates (Bottom) for Multi-Level Model of Video Fixations During Shape Word Interest Period in Expt. 1 (Observations = 3,200)

Fixed effect	Coefficient	SE	<i>t</i>	<i>p</i>
Intercept	-0.75	0.09	-8.69	< .001 ***
Contrastive accenting	-0.01	0.03	-0.09	.93
Beat gesture	0.02	0.03	0.57	.57
Contrast type	-0.09	0.03	-3.46	< .001 ***
Contrastive accenting x beat gesture	-0.02	0.05	-0.37	.71
Contrastive accenting x contrast type	0.26	0.34	0.76	.45
Beat gesture x contrast type	-0.01	0.05	-0.21	.83
Contrastive accenting x beat gesture x contrast type	0.01	0.11	0.10	.92
Gesture orientation	0.03	0.03	1.00	.32
Object side	-0.01	0.03	-0.33	.74
Gesture orientation x object side	0.01	0.06	0.08	.94

Random effect	s^2
Trial	0.13
Participant	0.53

Note:

 $p < .001$

Table 7

Fixed Effect (Top) and Variance Estimates (Bottom) for Multi-Level Model of Target Fixations During Trial Onset Interest Period in Expt. 1 (Observations = 3,206)

Fixed effect	Coefficient	SE	t	p
Intercept	-0.41	0.05	-7.66	< .001 ***
Contrastive accenting	-0.16	0.11	-1.52	.13
Beat gesture	0.14	0.06	2.31	.02*
Contrast type	0.01	0.04	0.31	.76
Trial	-0.03	0.02	-1.41	.16
Contrastive accenting x beat gesture	0.04	0.12	0.36	.72
Contrastive accenting x contrast type	0.28	0.20	1.41	.16
Contrastive accenting x trial	0.01	0.01	1.45	.15
Beat gesture x contrast type	-0.14	0.08	-1.75	.08 [†]
Beat gesture x trial	-0.01	0.01	-1.93	.05 [†]
Contrast type x trial	-0.01	0.01	-1.52	.13
Contrastive accenting x beat gesture x contrast type	-0.19	0.17	-1.11	.27
Contrastive accenting x beat gesture x trial	-0.01	0.01	-0.54	.59
Contrastive accenting x contrast type x trial	-0.01	0.01	-1.45	.15
Beat gesture x contrast type x trial	0.01	0.01	1.84	.07 [†]
Contrastive accenting x beat gesture x contrast type x trial	0.01	0.01	0.73	.47
Gesture orientation	-0.01	0.01	-0.07	.94
Object side	-0.01	0.02	-0.63	.53
Gesture orientation x object side	0.01	0.04	0.14	0.89
Random effect s^2				
Participant				0.29
Participant x beat gesture				0.09

* $p < .05$;

*** $p < .001$;

[†] $p < .10$

Table 8

Fixed Effect (Top) and Variance Estimates (Bottom) for Multi-Level Model of Competitor Fixations During Trial Onset Interest Period in Expt. 1 (Observations = 3,206)

Fixed effect	Coefficient	SE	<i>t</i>	<i>p</i>
Intercept	-0.42	0.05	-7.77	< .001***
Contrastive accenting	-0.13	0.11	-1.24	.22
Beat gesture	0.10	0.06	1.71	.09 [†]
Contrast type	-0.01	0.05	0.04	.97
Trial	-0.01	0.06	0.49	.63
Contrastive accenting x beat gesture	0.08	0.12	0.65	.51
Contrastive accenting x contrast type	0.19	0.20	0.98	.33
Contrastive accenting x trial	0.01	0.01	0.72	.47
Beat gesture x contrast type	-0.13	0.08	-1.63	.10
Beat gesture x trial	-0.01	0.01	-1.24	.21
Contrast type x trial	-0.01	0.01	-0.01	.99
Contrastive accenting x beat gesture x contrast type	-0.05	0.02	-0.31	.76
Contrastive accenting x beat gesture x trial	-0.01	0.01	-1.21	.23
Contrastive accenting x contrast type x trial	-0.01	0.01	-0.37	.71
Beat gesture x contrast type x trial	0.01	0.01	1.97	.05 [†]
Contrastive accenting x beat gesture x contrast type x trial	0.01	0.01	0.35	.72
Gesture orientation	-0.02	0.02	-1.15	.25
Object side	-0.01	0.02	-0.57	.57
Gesture orientation x object side	0.03	0.04	0.68	.49

Random effect	<i>s</i> ²
Participant	0.29
Participant x contrast type	0.13

Note:

p < .001;

[†]
p < .10

Table 9

Fixed Effect (Top) and Variance Estimates (Bottom) for Multi-Level Model of Target Fixations During Color Word Interest Period in Expt. 1 (Observations = 3,212)

Fixed effect	Coefficient	SE	<i>t</i>	<i>p</i>
Intercept	-0.31	0.05	-6.56	< .001 ***
Contrastive accenting	0.03	0.09	0.27	.79
Beat gesture	0.01	0.06	0.23	.82
Contrast type	-0.02	0.04	-0.56	.57
Trial	-0.01	0.02	-0.86	.39
Contrastive accenting x beat gesture	-0.12	0.12	-1.01	.31
Contrastive accenting x contrast type	0.07	0.17	0.42	.68
Contrastive accenting x trial	-0.01	0.01	-0.42	.67
Beat gesture x contrast type	0.30	0.07	3.98	< .001 ***
Beat gesture x trial	0.01	0.01	0.14	.89
Contrast type x trial	-0.01	0.01	-0.10	.92
Contrastive accenting x beat gesture x contrast type	0.07	0.18	0.43	.67
Contrastive accenting x beat gesture x trial	0.01	0.01	1.70	.09 †
Contrastive accenting x contrast type x trial	-0.01	0.01	-0.01	.99
Beat gesture x contrast type x trial	-0.01	0.01	-2.96	.003 **
Contrastive accenting x beat gesture x contrast type x trial	-0.01	0.01	-0.95	.34
Gesture orientation	0.01	0.01	1.58	.11
Object side	0.02	0.02	1.09	.28
Gesture orientation x object side	0.03	0.03	0.89	.38
<hr/>				
Random effect	<i>s</i> ²			
Participant	0.25			
Participant x beat gesture	0.16			

Note:

**
p < .01;

p < .001;

†
p < .10

Table 10

Tukey HSD-Corrected Post-Hoc Tests for Beat Gesture x Contrast Type Interaction in Model of Target Fixations During Color Word Interest Period in Expt. 1

Comparison	Estimate	SE	z-ratio	p
No beat, color vs. no beat, color + shape	0.07	0.02	3.11	.01*
No beat, color vs. beat, color	-0.02	0.03	-0.58	.94
No beat, color + shape vs. beat, color + shape	-0.12	0.03	-3.43	.003*
Beat, color vs. beat, color + shape	-0.03	0.02	-1.04	.73

Note:

*
p < .05

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 11

Fixed Effect (Top) and Variance Estimates (Bottom) for Multi-Level Model of Competitor Fixations During Color Word Interest Period in Expt. 1 (Observations = 3,212)

Fixed effect	Coefficient	SE	<i>t</i>	<i>p</i>
Intercept	-0.36	0.05	-7.68	< .001***
Contrastive accenting	0.03	0.09	0.34	.74
Beat gesture	-0.05	0.06	-0.80	.42
Contrast type	0.06	0.04	1.66	.10 [†]
Trial	0.01	0.02	0.36	.72
Contrastive accenting x beat gesture	-0.04	0.11	-0.33	.74
Contrastive accenting x contrast type	0.05	0.17	0.30	.77
Contrastive accenting x trial	-0.01	0.01	-0.67	.50
Beat gesture x contrast type	0.19	0.07	2.60	.009**
Beat gesture x trial	0.01	0.01	1.26	.21
Contrast type x trial	-0.01	0.01	-0.32	.75
Contrastive accenting x beat gesture x contrast type	-0.14	0.17	-0.84	.40
Contrastive accenting x beat gesture x trial	0.01	0.01	0.62	.54
Contrastive accenting x contrast type x trial	0.01	0.01	1.01	.31
Beat gesture x contrast type x trial	-0.01	0.01	1.44	.15
Contrastive accenting x beat gesture x contrast type x trial	0.01	0.01	0.63	.53
Gesture orientation	0.01	0.01	1.81	.07 [†]
Object side	0.01	0.02	0.01	.99
Gesture orientation x object side	-0.06	0.01	-1.71	.09 [†]
<hr/>				
Random effect	<i>s</i> ²			
Participant	0.25			
Participant x beat gesture	0.16			

Note:

**
p < .01;

p < .001;

[†]
p < .10

Table 12

Tukey HSD-Corrected Post-Hoc Tests for Beat Gesture x Contrast Type Interaction in Model of Competitor Fixations During Color Word Interest Period in Expt. 1

Comparison	Estimate	SE	z-ratio	p
No beat, color vs. no beat, color + shape	-0.01	0.02	-0.09	.99
No beat, color vs. beat, color	-0.01	0.03	-0.41	.98
No beat, color + shape vs. beat, color + shape	-0.11	0.03	-3.35	.005**
Beat, color vs. beat, color + shape	-0.10	0.02	-4.07	< .001***

Note:

**
p < .01;

p < .001

Table 13

Fixed Effect (Top) and Variance Estimates (Bottom) for Multi-Level Model of Target Fixations During Shape Word Interest Period in Expt. 1 (Observations = 3,212)

Fixed effect	Coefficient	SE	<i>t</i>	<i>P</i>
Intercept	-0.06	0.06	-0.97	.33
Contrastive accenting	0.04	0.08	0.51	.61
Beat gesture	0.02	0.07	0.29	.77
Contrast type	-0.54	0.07	-7.01	< .001***
Trial	0.04	0.03	1.36	.17
Contrastive accenting x beat gesture	-0.28	0.14	-1.94	.05 [†]
Contrastive accenting x contrast type	0.15	0.23	0.63	.53
Contrastive accenting x trial	-0.01	0.01	-0.43	.66
Beat gesture x contrast type	-0.08	0.14	-0.54	.59
Beat gesture x trial	-0.01	0.01	-0.25	.80
Contrast type x trial	0.01	0.01	3.01	.003**
Contrastive accent x beat gesture x contrast type	0.50	0.29	1.75	.08 [†]
Contrastive accenting x beat gesture x trial	0.01	0.01	2.48	.01*
Contrastive accenting x contrast type x trial	0.01	0.01	0.81	.42
Beat gesture x contrast type x trial	0.01	0.01	2.00	.05*
Contrastive accenting x beat gesture x contrast type x trial	-0.01	0.01	-0.90	.37
Gesture orientation	-0.01	0.01	-7.26	< .001***
Object side	-0.02	0.03	-0.69	.49
Gesture orientation x object side	0.49	0.06	7.72	< .001***

Random effect	<i>s</i> ²
Participant	0.29
Participant x contrast type	0.17

Note:

* $p < .05$;

** $p < .01$;

*** $p < .001$;

[†] $p < .10$

Table 14

Fixed Effect (Top) and Variance Estimates (Bottom) for Multi-Level Model of Competitor Fixations During Shape Word Interest Period in Expt. 1 (Observations = 3,212)

Fixed effect	Coefficient	SE	<i>t</i>	<i>p</i>
Intercept	-1.20	0.03	-34.47	<.001***
Contrastive accenting	0.06	0.05	1.25	.21
Beat gesture	-0.02	0.04	-0.42	.67
Contrast type	0.19	0.05	3.77	<.001***
Trial	0.01	0.01	2.40	.02*
Contrastive accenting x beat gesture	0.09	0.08	1.06	.29
Contrastive accenting x contrast type	0.33	0.14	2.36	.02*
Contrastive accenting x trial	0.01	0.01	0.23	.81
Beat gesture x contrast type	-0.01	0.08	-0.16	.87
Beat gesture x trial	-0.01	0.01	-0.30	.76
Contrast type x trial	-0.01	0.01	-0.33	.74
Contrastive accenting x beat gesture x contrast type	-0.20	0.17	-1.17	.24
Contrastive accenting x beat gesture x trial	-0.01	0.01	-0.73	.47
Contrastive accenting x contrast type x trial	-0.01	0.01	-0.48	.63
Beat gesture x contrast type x trial	0.01	0.01	0.37	.71
Contrastive accenting x beat gesture x contrast type x trial	0.01	0.01	1.18	.24
Gesture orientation	0.01	0.02	0.44	.66
Object side	0.01	0.02	0.61	.54
Gesture orientation x object side	-0.09	0.04	-2.36	.02*

Random effect	s^2
Participant	0.17
Participant x contrast type	0.17

Note:

* $p < .05$;

*** $p < .001$

Table 15

Tukey HSD-Corrected Post-Hoc Tests for Contrastive Accenting x Contrast Type Interaction in Model of Competitor Fixations During Shape Word Interest Period in Expt. 1

Comparison	Estimate	SE	z-ratio	p
Non-contrastive, color vs. non-contrastive, color + shape	-0.03	0.07	-0.48	.96
Non-contrastive, color vs. contrastive, color	0.07	0.06	1.22	.62
Non-contrastive, color + shape vs. contrastive, color + shap	-0.22	0.07	-2.94	.02*
Contrastive, color vs. contrastive, color + shape	-0.32	0.07	-4.77	< .001***

Note:

*
p < .05;

p < .001

Table 16

Proportion of Fixations to Each Interest Area by Interest Period in Expt. 2

Interest period	Video			Initial	Target	Competitor	Distractor
	Beat	No beat	Overall				
Trial onset	0.82	0.80	0.81	0.12	0.02	0.03	0.03
Color word	0.88	0.81	0.84	0.06	0.03	0.03	0.03
Shape word	0.44	0.41	0.43	0.03	0.36	0.08	0.10

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 17

Fixed Effect (Top) and Variance Estimates (Bottom) for Multi-Level Model of Video Fixations During Trial Onset Interest Period in Expt. 2 (Observations = 3,115)

Fixed effect	Coefficient	SE	<i>t</i>	<i>P</i>
Intercept	-0.14	0.05	-2.88	.005**
Contrastive accenting	-0.01	0.03	-0.35	.73
Beat gesture	-0.01	0.03	-0.34	.73
Contrast type	-0.01	0.02	-0.33	.74
Contrastive accenting x beat gesture	-0.08	0.06	-1.34	.18
Contrastive accenting x contrast type	-0.01	0.04	-0.16	.88
Beat gesture x contrast type	-0.02	0.04	-0.53	.60
Contrastive accenting x beat gesture x contrast type	0.07	0.09	0.82	.41
Gesture orientation	-0.02	0.03	-0.51	.61
Object side	0.02	0.03	0.59	.56
Gesture orientation x object side	-0.01	0.05	-0.02	.98

Random effect	<i>s</i> ²
Trial	0.13
Participant	0.25

**
p < .01

Table 18

Fixed Effect (Top) and Variance Estimates (Bottom) for Multi-Level Model of Video Fixations During Color Word Interest Period in Expt. 2 (Observations = 3,115)

Fixed effect	Coefficient	SE	t	p
Intercept	-0.90	0.05	-16.53	< .001***
Contrastive accenting	0.05	0.04	1.32	.19
Beat gesture	-0.02	0.04	-0.60	.54
Contrast type	0.11	0.03	3.94	< .001***
Contrastive accenting x beat gesture	0.45	0.08	5.61	< .001***
Contrastive accenting x contrast type	0.01	0.06	0.12	.90
Beat gesture x contrast type	0.08	0.06	1.38	.17
Contrastive accenting x beat gesture x contrast type	-0.34	0.11	-2.96	.003**
Gesture orientation	0.08	0.04	1.91	.06*
Object side	0.28	0.04	6.52	< .001***
Gesture orientation x object side	-0.31	0.06	-5.07	< .001***
Random effect σ^2				
Trial	0.23			
Participant	0.22			

Table 19

Fixed Effect (Top) and Variance Estimates (Bottom) for Multi-Level Model of Video Fixations During Shape Word Interest Period in Expt. 2 (Observations = 3,115)

Fixed effect	Coefficient	SE	t	p
Intercept	0.13	0.08	1.74	.09 [†]
Contrastive accenting	0.05	0.03	1.54	.12
Beat gesture	0.04	0.03	1.26	.21
Contrast type	-0.06	0.03	-1.92	.05 [†]
Contrastive accenting x beat gesture	-0.03	0.06	-0.46	.64
Contrastive accenting x contrast type	0.02	0.06	0.37	.71
Beat gesture x contrast type	-0.14	0.06	-2.30	.02 [*]
Contrastive accenting x beat gesture x contrast type	0.04	0.12	3.21	.001 ^{**}
Gesture orientation	-0.07	0.04	-0.15	.88
Object side	-0.04	0.05	-0.83	.41
Gesture orientation x object side	0.02	0.06	0.36	.72
Random effect s^2				
Trial	0.17			
Participant	0.42			

* $p < .05$;

** $p < .01$;

[†] $p < .10$

Table 20

Fixed Effect (Top) and Variance Estimates (Bottom) for Multi-Level Model of Target Fixations During Trial Onset Interest Period in Expt. 2 (Observations = 3,115)

Fixed effect	Coefficient	SE	<i>t</i>	<i>p</i>
Intercept	-0.02	0.01	2.17	.03*
Contrastive accenting	-0.01	0.02	-0.05	.96
Beat gesture	-0.02	0.02	-0.82	.42
Contrast type	-0.02	0.01	-1.14	.26
Trial	-0.01	0.01	-6.14	< .001***
Contrastive accenting x beat gesture	-0.01	0.04	-0.02	.98
Contrastive accenting x contrast type	0.02	0.03	0.66	.51
Contrastive accenting x trial	0.01	0.01	0.67	.50
Beat gesture x contrast type	0.03	0.03	1.23	.22
Beat gesture x trial	0.01	0.01	1.25	.21
Contrast type x trial	0.01	0.01	1.36	.17
Contrastive accenting x beat gesture x contrast type	0.01	0.06	0.10	.92
Contrastive accenting x beat gesture x trial	-0.01	0.01	-0.18	.85
Contrastive accenting x contrast type x trial	-0.01	0.01	-0.69	.49
Beat gesture x contrast type x trial	-0.01	0.01	-2.34	.02*
Contrastive accenting x beat gesture x contrast type x trial	0.01	0.01	0.22	.83
Gesture orientation	-0.01	0.01	-0.56	.58
Object side	0.01	0.01	1.36	.17
Gesture orientation x object side	0.02	0.01	-1.91	.06 [†]

Random effect	<i>s</i> ²
Participant	0.03
Participant x beat gesture	0.05

* $p < .05$;

*** $p < .001$;

[†] $p < .10$

Table 21

Fixed Effect (Top) and Variance Estimates (Bottom) for Multi-Level Model of Competitor Fixations During Trial Onset Interest Period in Expt. 2 (Observations = 3,115)

Fixed effect	Coefficient	SE	t	p
Intercept	0.02	0.01	1.93	.05 [†]
Contrastive accenting	0.01	0.02	0.35	.73
Beat gesture	-0.01	0.02	-0.52	.60
Contrast type	-0.02	0.01	-1.38	.17
Trial	-0.01	0.01	-5.48	< .001***
Contrastive accenting x beat gesture	-0.02	0.04	-0.62	.53
Contrastive accenting x contrast type	0.01	0.03	0.25	.80
Contrastive accenting x trial	0.01	0.01	0.06	.95
Beat gesture x contrast type	0.02	0.03	0.86	.39
Beat gesture x trial	0.01	0.01	0.73	.47
Contrast type x trial	0.01	0.01	1.84	.07 [†]
Contrastive accenting x beat gesture x contrast type	0.03	0.05	0.63	.53
Contrastive accenting x beat gesture x trial	0.01	0.01	1.11	.27
Contrastive accenting x contrast type x trial	-0.01	0.01	-0.09	.93
Beat gesture x contrast type x trial	-0.01	0.01	-1.86	.06 [†]
Contrastive accenting x beat gesture x contrast type x trial	-0.01	0.01	-0.82	.41
Gesture orientation	-0.01	0.01	-2.15	.03*
Object side	0.01	0.01	0.58	.56
Gesture orientation x object side	-0.02	0.01	-2.03	.04*

Random effect	s ²
Participant	0.02
Participant x beat gesture	0.05

* $p < .05$;
 *** $p < .001$;
[†] $p < .10$

Table 22

Fixed Effect (Top) and Variance Estimates (Bottom) for Multi-Level Model of Target Fixations During Color Word Interest Period in Expt. 2 (Observations = 3,115)

Fixed effect	Coefficient	SE	<i>t</i>	<i>p</i>
Intercept	-0.25	0.04	-7.04	< .001 ***
Contrastive accenting	0.03	0.05	0.62	.53
Beat gesture	-0.03	0.05	-0.68	.50
Contrast type	0.05	0.03	1.46	.14
Trial	0.01	0.01	1.17	.24
Contrastive accenting x beat gesture	-0.04	0.10	-0.42	.67
Contrastive accenting x contrast type	-0.22	0.07	-3.16	.002 **
Contrastive accenting x trial	-0.01	0.01	-0.40	.69
Beat gesture x contrast type	-0.11	0.07	-1.63	.10
Beat gesture x trial	0.01	0.01	-0.82	.41
Contrast type x trial	-0.01	0.01	-1.40	.16
Contrastive accenting x beat gesture x contrast type	0.24	0.14	1.79	.07 †
Contrastive accenting x beat gesture x trial	0.01	0.01	0.08	.94
Contrastive accenting x contrast type x trial	0.01	0.01	3.15	.002 **
Beat gesture x contrast type x trial	0.01	0.01	1.25	.21
Contrastive accenting x beat gesture x contrast type x trial	-0.01	0.01	-1.57	.12
Gesture orientation	-0.01	0.02	-0.41	.68
Object side	0.01	0.02	0.83	.41
Gesture orientation x object side	0.02	0.03	0.53	.61

Random effect	<i>s</i> ²
Participant	0.17
Participant x beat gesture	0.11

Note:

**
p < .01;

p < .001;

†
p < .10

Table 23

Fixed Effect (Top) and Variance Estimates (Bottom) for Multi-Level Model of Competitor Fixations During Color Word Interest Period in Expt. 2 (Observations = 3,115)

Fixed effect	Coefficient	SE	<i>t</i>	<i>p</i>
Intercept	-0.24	0.03	-6.95	< .001***
Contrastive accenting	-0.02	0.05	-0.44	.66
Beat gesture	-0.07	0.05	-1.46	.14
Contrast type	0.04	0.04	1.21	.23
Trial	0.01	0.01	1.00	.32
Contrastive accenting x beat gesture	-0.09	0.10	-0.88	.38
Contrastive accenting x contrast type	-0.16	0.07	-2.31	.02*
Contrastive accenting x trial	0.01	0.01	0.53	.60
Beat gesture x contrast type	-0.04	0.07	-0.64	.52
Beat gesture x trial	0.01	0.01	1.48	.14
Contrast type x trial	-0.01	0.01	-1.14	.26
Contrastive accenting x beat gesture x contrast type	0.21	0.14	1.52	.13
Contrastive accenting x beat gesture x trial	0.01	0.01	0.67	.50
Contrastive accenting x contrast type x trial	0.01	0.01	2.16	.03*
Beat gesture x contrast type x trial	0.01	0.01	0.34	.73
Contrastive accenting x beat gesture x contrast type x trial	-0.01	0.01	-1.45	.15
Gesture orientation	-0.20	0.02	-1.24	.21
Object side	0.03	0.02	-1.98	.05 [†]
Gesture orientation x object side	-0.02	0.03	-0.62	.54

Random effect	<i>s</i> ²
Participant	0.16
Participant x contrast type	0.06

Note:

* $p < .05$;

*** $p < .001$;

[†] $p < .10$

Table 24

Fixed Effect (Top) and Variance Estimates (Bottom) for Multi-Level Model of Target Fixations During Shape Word Interest Period in Expt. 2 (Observations = 3,115)

Fixed effect	Coefficient	SE	<i>t</i>	<i>p</i>
Intercept	-0.29	0.04	-7.99	< .001***
Contrastive accenting	-0.03	0.04	-0.81	.42
Beat gesture	0.02	0.04	0.43	.67
Contrast type	0.05	0.04	1.32	.19
Trial	0.01	0.01	1.58	.12
Contrastive accenting x beat gesture	0.01	0.07	0.14	.89
Contrastive accenting x contrast type	-0.01	0.07	-1.15	.25
Contrastive accenting x trial	0.01	0.01	1.56	.12
Beat gesture x contrast type	-0.15	0.07	-2.12	.03*
Beat gesture x trial	-0.01	0.01	-0.02	.99
Contrast type x trial	-0.01	0.01	-1.79	.07 [†]
Contrastive accenting x beat gesture x contrast type	-0.03	0.14	-0.24	.81
Contrastive accenting x beat gesture x trial	-0.01	0.01	-0.12	.90
Contrastive accenting x contrast type x trial	0.01	0.01	1.51	.13
Beat gesture x contrast type x trial	0.01	0.01	1.10	.27
Contrastive accenting x beat gesture x contrast type x trial	0.01	0.01	1.01	.31
Gesture orientation	0.01	0.02	0.07	.94
Object side	-0.02	0.02	-0.94	.35
Gesture orientation x object side	-0.01	0.03	-0.21	.84

Random effect	<i>s</i> ²
Participant	0.20
Participant x beat gesture	0.06

Note:

*
p < .05;

p < .001;

[†]
p < .10

Table 26

Fixed Effect (Top) and Variance Estimates (Bottom) for Multi-Level Model of Competitor Fixations During Shape Word Interest Period in Expt. 2 (Observations = 3,115)

Fixed effect	Coefficient	SE	<i>t</i>	<i>p</i>
Intercept	-0.28	0.03	-8.12	< .001 ***
Contrastive accenting	0.01	0.04	0.16	.87
Beat gesture	0.02	0.04	0.47	.64
Contrast type	0.02	0.04	0.68	.50
Trial	0.01	0.01	1.70	.09 [†]
Contrastive accenting x beat gesture	0.06	0.07	0.86	.39
Contrastive accenting x contrast type	-0.07	0.07	-0.98	.33
Contrastive accenting x trial	0.01	0.01	0.11	.91
Beat gesture x contrast type	-0.20	0.07	-2.89	.004 **
Beat gesture x trial	-0.01	0.01	-0.26	.79
Contrast type x trial	-0.01	0.01	-0.90	.37
Contrastive accenting x beat gesture x contrast type	-0.01	0.14	-0.09	.93
Contrastive accenting x beat gesture x trial	-0.01	0.01	-1.49	.14
Contrastive accenting x contrast type x trial	0.01	0.01	0.82	.41
Beat gesture x contrast type x trial	0.01	0.01	2.49	.01*
Contrastive accenting x beat gesture x contrast type x trial	0.01	0.01	0.98	.33
Gesture orientation	0.01	0.02	0.32	.75
Object side	0.01	0.02	0.83	.41
Gesture orientation x object side	-0.02	0.03	-0.55	.58
Random effect				
Participant	0.19			
Participant x beat gesture	0.10			

Note:

* $p < .05$;

** $p < .01$;

*** $p < .001$;

[†] $p < .10$