Distribution Agreement

_________________________ _________________________

Josephine Duan 04/17/13

The Sound of Color: Do Prosodic Cues Reflect Perceived Color and Brightness?

by

Josephine Duan

Adviser

Lynne Nygaard, PhD

Department of Psychology

____
Lynne Nygaard, PhD
Adviser

____

Laura Namy, PhD
Committee Member

____
Cheryl Crowley, PhD
Committee Member

2013

The Sound of Color: Do Prosodic Cues Reflect Perceived Color and Brightness?

By

Josephine Duan

Adviser
Lynne Nygaard, PhD

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Department of Psychology

2013

Abstract

The Sound of Color: Do Prosodic Cues Reflect Perceived Color and Brightness?

By Josephine Duan

Although prosody (the rhythm, intonation, and rate of speech) has been assumed to have no direct impact on the processing of meaning, recent studies have shown that language users may be sensitive to prosodic cues to meaning in spoken language, suggesting that prosody may impact how we produce and perceive semantic information in speech. The current study examined one potential example of non-arbitrary sound to meaning mapping that may be reflected in prosody, the correspondence between sound and color. Two experiments examined the extent to which speakers spontaneously produce prosodic correlates to meaning when describing color and brightness. In Experiment 1, participants produced verbal labels for different colors and brightness levels using actual color names. In Experiment 2, participants performed a similar task using novel words for color labels. Acoustic analyses in Experiment 1 showed that speakers' prosody reliably differed across color labels. Further, labels for cool colors (purple, blue, and green) were significantly higher in pitch and amplitude than those for warm colors (yellow, orange, and red). However, the same differences were not found in Experiment 2. Findings suggest a complex relationship between prosody and meaning, which may be attributed to different types of auditory and visual mappings.

Keywords: prosody, semantics, cross-modal correspondences, color perception, brightness perception

The Sound of Color: Do Prosodic Cues Reflect Perceived Color and Brightness?

By

Josephine Duan

Adviser
Lynne Nygaard, PhD

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Department of Psychology

2013

Acknowledgements

This research was funded by a Scholarly Inquiry and Research at Emory (SIRE) grant. I would like to thank my thesis advisor, Dr. Lynne Nygaard, and my graduate mentor, Christina Tzeng, for their encouragement and support throughout all stages of this research. I am also thankful to my committee members, Dr. Laura Namy and Dr. Cheryl Crowley, for their insight and guidance on this project. I am also grateful to the members of the Speech and Language Perception Lab and Language Group for their thoughtful insights into experimental design and results.

Table of Contents

The Sound of Color: Do Prosodic Cues Reflect Perceived Color and Brightness?

Human communication occurs in many forms. People communicate nonverbally with their bodies, such as with a firm handshake, a gaze, or a stare. They also communicate verbally, primarily in the form of spoken language. Spoken language is comprised of linguistic elements, such as words and phrases, and extra-linguistic properties, such as the tone of voice. Traditionally, linguists have maintained that the semantic meaning of words and sentences is conveyed by the linguistic rather than extra-linguistic elements of spoken language. However, extra-linguistic elements also constitute an important part of the communicative power of spoken language. In particular, prosody, or the rhythm, intonation and rate of speech, serves a variety of communicative functions in speech. Acoustically, prosody is measured in fundamental frequency (pitch), amplitude (loudness), and relative duration. The current study sought to further investigate the functionality of prosody and examine how humans incorporate prosodic cues in their everyday speech.

Previous research on the functionality of prosody has mainly focused on how prosodic cues aid in the disambiguation of linguistic structure. Studies have found that prosodic cues such as stress patterns help listeners to identify word and phrase boundaries and facilitates syntax acquisition (Cutler, 1997). For example, Jusczyk et al. (1992) found that very young infants used prosody to determine phrase boundaries in sentence-length utterances. In their study, both nine-month-old infants and adults were able to parse English phrases even when phonetic information was removed. Acoustic analyses revealed that changes in pitch contour and syllable duration might have been used to identify structural boundaries. For example, major phrasal boundaries were preceded by declines in fundamental frequency and longer syllable durations.

Other prosodic markers in speech such as pause location also influence how listeners interpret the meaning of phrases (Clifton, Carlson, & Frazier, 2002). In Clifton et al.'s study (2002), participants listened to different versions of phrases (e.g., "old men and women with very large houses") with a pause placed either after "old men" or after "old men and women", and were asked to determine who has the large houses. Participants' answers varied as the speaker altered the prosodic boundaries in her speech. For example, when a pause was placed after the words "old men", participants were more likely to associate the ownership of the large houses to only the women than to both the old men and women. Thus, prosodic cues and markers in speech assist listeners in the disambiguation of syntax.

Additionally, research has also shown that prosody plays an important role in reflecting the emotional state, intentionality, and attitude of the speaker (e.g., Ofuka, McKeown, Waterman, & Roach, 2000; Scherer, 1994; Spence & Moore, 2003). As stated in Scherer's (1994) review, previous studies have consistently found that certain vocal cues are correlated with specific emotional states. For example, emotions such as joy and elation are significantly correlated with higher mean fundamental frequency, mean intensity, and speech rate (Scherer, 1994). On the other hand, emotions such as sadness and dejection are correlated with lower mean fundamental frequency, mean intensity and speech rate (Scherer, 1994). Further, studies have found that listeners can identify the emotional state or attitude of the speakers based on the speakers' variations in pitch, amplitude, speech rate, and other acoustic properties (e.g., Bachorowski, 1999; Ofuka et al., 2000). In Ofuka et al.'s study (2000), researchers found that the prosody of the utterances impacted how listeners perceived the politeness level of Japanese speech. In particular, the duration of the final vowel and overall speech rate correlated to the perceived level of politeness.

Although previous research has acknowledged the functional significance of prosody in human communication, researchers have mainly focused on how prosody supports the comprehension of linguistic elements in speech and its ability to convey nonlinguistic properties, such as speaker intentionality or emotion. Researchers have long assumed prosody to have no direct relationship to semantic meaning. Instead, linguistic units, such as words and sentences, have been traditionally assumed to convey meaning, and prosody was thought to have facilitated in communication only by aiding in the parsing of syllables and word boundaries or as a reflection of the emotional state of the speaker. Language is often viewed as a symbolic system in which information is conveyed through symbols whose forms and their referent meanings are arbitrarily related (Hockett, 1960; deSaussure, 1959). Thus, researchers have assumed that there is no systematic relationship between the sound and the meaning of a word.

However, researchers have found that prosody exhibits non-arbitrary sound-to-meaning relationships, in that acoustic properties of speech may directly convey semantic information. Specifically, studies have shown that prosodic cues, such as an emotional tone of voice, may interact with how people comprehend the meaning of emotion words (e.g., Nygaard & Lunders, 2002; Wurm, Vakoch, Strasser, Calin-Jageman, & Ross, 2001). For example, in Nygaard and Lunders (2002), participants heard emotional homophones such as *die/dye*, in which one of the words in the pair had a positive or negative meaning (e.g., *die*), while the other had a neutral meaning (e.g., *dye*). Participants listened to words that were produced in a positive, negative, or neutral tone of voice and were asked to transcribe the words they heard. Results indicated that when the tone of voice matched the emotional meaning of the homophone, participants were more likely to transcribe the emotional meaning than the neutral one. For example, if the word *die/dye* was spoken with a sad tone of voice, participants were more likely to transcribe the word as *die* than as *dye*. Thus, these findings are inconsistent with previous assumptions, which claim that prosody has no direct impact on the linguistic interpretation of speech. Instead, these results suggest that prosody may provide cues to word meaning.

More recently, researchers have also found evidence for the connection between prosody and word meaning for words that are not associated with emotion. Nygaard, Herold, and Namy (2009) found that when speakers were asked to produce auditory labels for novel words, they reliably produced acoustic profiles that are similar but slightly different for meanings that share similar dimensions (e.g., big and tall), suggesting the incorporation of a conceptual difference between dimensions such as tall and big in the acoustic properties of speech. One example is that labels reliably associated with a big picture were lower in amplitude and longer in duration than labels associated with a small picture. However, when compared to those of tall pictures, labels associated with big pictures were similar in amplitude and duration but varied in fundamental frequency. Participants also use prosody to disambiguate the meaning of novel dimensional adjectives. In the same Nygaard et al. (2009) study, participants heard the novel words described above and then chose which of the two pictures (one representing the word's meaning, and the other, its antonym) presented on the computer screen best represented the meaning of the word they heard. Results showed that participants were able to infer the meaning of the novel word more accurately when the speaker's prosody and the word meaning choices matched than when they mismatched. This suggests that participants were able to interpret the meaning of the novel word based solely on the prosodic cues in the speaker's voice.

In fact, Herold, Nygaard, and Namy (2012) found that prosody plays an important role in facilitating children's word learning. Mothers were asked to label pictures with dimensional English adjectives (e.g., big, small, happy, sad, hot, cold) while reading a picture book to their children. Acoustic analyses of recorded utterances revealed that mothers significantly altered their amplitude and duration for target words and sentences to distinguish between antonyms and dimensions. Thus, similar to the findings in Nygaard et al. (2009), participants recruited prosodic cues in their infant-directed speech (IDS; a form of speech often characterized by heightened vocal pitches and elongated syllables) to convey word meaning. Furthermore, not only does prosody aid in speech segmentation and emotion signaling, it also serves to convey referential information about word meaning, facilitating the process of word learning in children.

Additionally, researchers have found that speakers naturally incorporate prosodic cues in their speech to convey visuo-spatial information (Shintel, Nusbaum, & Okrent, 2006). Shintel et al. (2006) conducted a series of three experiments to examine the role of prosody in conveying visuo-spatial information about referent objects. In their first experiment, participants were presented with an animation of a dot moving either upward or downward. Participants were instructed to describe the direction of movement of the dot using the sentence “It is going up” or “It is going down”. Results indicated that participants spontaneously produced systematic prosodic cues to describe movements of a dot even when such movement could easily be described by the action term itself. For example, despite the fact that the words *up* and *down* have clear spatial meanings, participants’ utterances still varied in fundamental frequency (e.g., higher pitch for “up” vs. “down”) as they described the two different scenarios. This suggests a cross-modal correspondence between the auditory and the visual domains, in that people perceive an association between the level of fundamental frequency and visuo-spatial information such as vertical location. Furthermore, in the subsequent two experiments, Shintel et al. (2006) found that listeners were able to identify which visual event a speaker was describing exclusively from changes in that speaker’s prosody. For example, participants were able to infer the rate of a moving dot solely based on the speaking rate of the speaker. Thus, these findings suggest that prosodic properties of speech exhibit a non-arbitrary relationship between sound and meaning and can directly provide semantic-referential information.

Researchers have argued that the non-arbitrary relationship between sound and meaning exhibited by the prosodic properties of speech may be due to the grounding of language in perceptual simulations of described events (Barsalou, 1999). Prosodic cues may convey information about visual properties of a referent as a result of the general cross-modal nature of our perceptual systems (Barsalou, 1999). Thus, language users may recruit prosody to convey meaning in spoken language because acoustic properties such as pitch are systematically related to specific visual properties, such as spatial height (e.g., Shintel et al., 2006; Shintel and Nusbaum, 2007). When producing speech, speakers may simulate the experience of a given word or phrase and spontaneously reflect it in their acoustic structure of speech. Indeed, research on non-speech tones has shown that perceivers consistently associate auditory sensory properties, such as pitch and amplitude, with particular visual sensory attributes, such as size and brightness levels (e.g., Marks, 1982; Marks, 1987). Marks' (1982) study examined synaesthetic metaphors that incorporated words or phrases typically used to describe experiences in one modality to describe experiences in another modality (e.g., dawn comes up like thunder). Across a series of four experiments, participants were asked to rate these expressions on scales of loudness, pitch, and brightness (Marks, 1982). Results indicated that greater levels of brightness were associated with louder amplitude and higher pitch when brightness referred to luminosity (whether or not the perceived object was glowing). When brightness corresponded to surface reflectance (whether or not the perceived object was highly reflective), greater brightness was only associated with higher pitch but not louder amplitude (Marks, 1982). Thus, Marks (1982) suggested that people perceive cross-modal correspondences between visual and auditory domains and that these cross-modal correspondences may have occurred due to similar sensory perceptions across modalities.

Marks, Hammeal, and Bornstein (1987) further demonstrated that levels of visual brightness were associated with different levels of pitch, such that higher pitch sounds are associated with brighter colors, and lower pitch sounds are correlated to darker colors. Participants were asked to judge the color terms (e.g., yellow, red, green, blue, brown) in terms of pitch levels. Results indicated that both children and adults rated yellow as higher pitched and smaller sized than brown. Thus, participants reliably associated different levels of fundamental frequency or pitch with different colors, with brighter colors (e.g., yellow) associated with higher pitch sounds than darker colors (e.g., blue, brown). As Marks (1982) mentions, synaesthetic individuals often perceive sounds in terms of visual shapes and colors. For synaesthetic individuals, pitch is reliably associated with the perception of brightness levels (Marks, 1982). Interestingly, non-synaesthetes also exhibit similar cross-modal mappings, suggesting that perceptual properties across modalities may elicit similar sensory perceptions.

The current study sought to investigate language users' sensitivity to cross-modal correspondences between visual and auditory domains and whether speakers' utterances will reflect these underlying mappings. In particular, the current investigation examined prosodic correlates to color and brightness as a potential example of non-arbitrary auditory-visual correspondences in spoken language. Although previous studies on non-speech tones have suggested that listeners systematically associate acoustic properties with color, whether these relationships are evident in spoken words has yet to be explored. We examined whether speakers use prosody to differentiate between different levels of brightness within a color and whether or not prosodic correlates vary across different colors. Specifically, the current study examined the extent to which language users recruit different levels of pitch, amplitude, and duration when referring to visual representations of color brightness. In the current experiment, participants were asked to produce utterances such as “Can you get the _______ one?” for different colors and different brightness levels of colors. In Experiment 1, participants were asked to produce utterances using color labels (e.g., blue, green, orange, purple, red, yellow). In Experiment 2, participants were asked to produce utterances using novel word labels (e.g., *blicket, daxen, foppick, tillen, riffel, seebow*). Acoustic analyses were conducted on the recorded utterances to investigate whether speakers produce reliable acoustic properties across brightness levels and across colors.

If prosodic correlates to meaning rely on general cross-modal correspondences, language users should produce and incorporate prosodic cues that reflect the relationship between pitch and different levels of brightness of a color in their utterances. We hypothesized that when referring to lighter shades of a particular color, speakers would reliably produce labels with a higher pitch (e.g., light yellow would be correlated with a higher pitch utterance). When referring to darker shades of a particular color, participants would reliably produce labels with lower pitches (e.g., dark yellow would be correlated with a lower pitch utterance). In addition, in alignment with Marks et al.’s study (1987), we predicted that speakers would produce labels with a higher pitch for brighter colors (e.g., yellow) than darker colors (e.g., blue).

Experiment 1

The objective of Experiment 1 was to investigate the extent to which speakers spontaneously produce prosodic correlates to color brightness. Brightness in the current experiment is defined as the relative lightness or darkness of a particular color, in which black represents no brightness, whereas white represents full brightness. Participants completed a production task designed to elicit verbal labels for different brightness levels of various colors. We predicted that the prosodic realization of verbal labels for the lighter shades of a particular color would differ significantly from those produced for the darker shades of that color. Specifically, we predicted that lighter shades would be produced with higher mean fundamental frequencies than darker shades. We also anticipated observing differences in the production of prosody across the different colors. As shown in Marks et al. (1987), brighter colors such as yellow should be associated with higher pitch sounds than darker colors such as blue or brown.

Methods

Participants. Thirty-six Emory University undergraduate students participated in the experiment. Analyses were conducted with data from 20 participants (13 participants were excluded due to familiarity with a foreign language before age seven; three were excluded due to failure to record utterances on too many trials). All participants were monolingual American English speakers with no reported history of speech or hearing disorders. Participants were recruited from the Emory University Introductory Psychology subject pool and received course credit for participation.

Materials. Stimuli consisted of color spectrum displays of each of six colors; red, orange, yellow, green, blue, or purple. Each color spectrum included nine different shades of each color that varied only in brightness by RGB coordinates (See Appendix A for all color spectrums and coordinates). Each color shade was obtained from http://www.rapidtables.com/web/color/RGB_Color.htm using the color chart provided. The individual shades in yellow and purple color spectrums were adjusted due to high level of similarity between adjacent shades. Each spectrum was 9 x 1.85 inches, with the width and height of each shade equated. Three different versions of each color spectrum were created. In each version, an arrow was placed above the lightest, darkest, or intermediate shade to indicate for which shade participants should be producing verbal labels.

Procedure. Participants were asked to provide spoken labels for light, intermediate, and dark shades of each of the six colors. The experiment consisted of three blocks. Each block consisted of 18 trials in which each version of the color spectrums (e.g., yellow spectrum with the arrow pointing at the lightest shade) was presented once. Trials within each block were presented in random order. For each trial, participants were instructed to describe the shade of the color as best as they could to an imaginary listener using *only* the color label provided above the spectrum. Participants were asked to respond to each specific color spectrum by using the sentence "Can you get the _____ one?" filling in the blank with the provided color label (e.g., purple). To ensure that participants would use the appropriate wording in their utterances, each participant was asked to complete a practice trial during which the experimenter presented a gray color spectrum and asked participants to respond as they would in the actual experiment.

At the beginning of each trial, a fixation cross was presented at the center of the display for 500ms. Next, a color spectrum was presented with an arrow above indicating either the darkest, lightest, or an intermediate shade of a color. The arrow indicated the shade for which the participants were producing verbal labels. A color label (e.g., purple) was placed above the arrow indicating the displayed color. Below the color spectrum was the sentence "Can you get the ______ one?". Each color spectrum remained on the display for 5s during which participants produced the target sentence with the appropriate color label (e.g., Can you get the purple one?"). Participants were instructed to initiate their responses quickly once the color spectrum was presented. Utterances were recorded in a sound-attenuated room using two audio-technica

ATR 20 microphones onto a Dell computer and segmented by trials using Eprime 2.0 (Psychology Software Tools, Inc). Sentence utterances were re-digitized at a 22.050 kHz sampling rate and amplitude normalized using PRAAT (Boersma & Weenink, 2012), a speech analysis software. To examine the acoustic features of the color labels, each color label was segmented from the sentence utterance and amplitude normalized.

Acoustic Analyses. For each sentence and word utterance, acoustic measures were obtained to determine if participants' productions varied as a function of brightness and color. Potential prosodic correlates to color brightness and meaning may not be limited to the single-word color labels in this task; thus, analyses were conducted on both the entire sentence utterance and separately on the color labels segmented from the sentence. Mean fundamental frequency (F_0), mean amplitude, and duration were measured using PRAAT. F_0 refers to the number of cycles per second in a periodic sound and corresponds to the perception of pitch. Amplitude reflects the overall energy of the utterance and corresponds to the perception of loudness. Duration is the overall length of the utterance.

Results and Discussion

Separate acoustic analyses were conducted on sentence level and single-word level utterances. To assess the overall performance collapsed across colors for F_0, amplitude, and duration, three (one for each dependent measure) one-way repeated-measures analysis of variance (ANOVA) were conducted with brightness level (Dark, Intermediate, Light) as a within-subjects factor on single-word level utterances. Results did not indicate a significant difference across brightness levels for any of the three dependent measures (F_0, F (2,38)= .95, p=.40, *partial* η^2=.05; Amplitude, F (2,38)= 1.60, p=.22, *partial* η^2=.08; Duration, F (2,38)=1.00, p=.38, *partial* η^2=.05), suggesting that speakers did not reliably produce prosodic correlates to brightness levels (see Fig. 1a, 2a, 3a). However, the difference between dark and light shades trended in the predicted direction, in that the mean F_0 for the lightest shades (*M*=166.54; *SD*=37.52) was higher than that for the darkest shades (*M*=164.57; *SD*=38.75; see Fig. 1a).

Three three-way repeated-measures ANOVAs with brightness level, color (purple, blue, green, yellow, orange, red) and block (Block 1, 2, 3) as within-subjects factors, revealed a main effect of color for all three dependent measures (F_0, F (5,95)= 2.65, p=.03, *partial* η^2=.03; Amplitude, F (5,95)=43.83, p<.001, *partial* η^2=.70; Duration, F (5,95)= 208.10, p< .001, *partial* η^2=.92). No main effect of brightness level (F_0, F (2,38)=.98, p=.39, *partial* η^2=.05; Amplitude, F (2,38)=1.60, p=.21, *partial* η^2=.08; Duration, F (2,38)=.77, p=.47, *partial* η^2=.04) or block (F_0, F (2,38)=.67, p=.52, *partial* η^2=.03; Amplitude, F (2,38)=.52, p=.60, *partial* η^2=.03; Duration, F (2,38)=2.44, p=.10, *partial* η^2=.11) was found. No significant interactions were found. Pairwise comparisons of colors were conducted on all three dependent measures and were adjusted for multiple comparisons. These analyses are reported below.

Pitch. Fig. 4a shows the average of mean F_0 values across colors at the word-level. Labels for blue (*M*=169.41; *SD*=40.75) were significantly higher pitched than those for orange (*M*=162.53; *SD*=36.36, *p*=.02), red (*M*=164.42; *SD*=39.01, *p*=.003), and yellow (*M*=162.45; *SD*=37.52, *p*=.01). Labels for green (*M*=167.79; *SD*=37.07) were significantly higher pitched than those for orange (*p*=.03) and yellow (*p*=.005). Labels for orange were significantly lower pitched than those for blue (*p*=.02), green (*p*=.03), and purple (*M*=168.04; *SD*=44.34, *p*=.04). Labels for purple were significantly higher pitched than those for orange (*p*=.04). Labels for red were significantly lower pitched than those for blue (*p*=.003). Labels for yellow were significantly lower pitched than those for blue (*p*=.01) and green (*p*=.005).

Amplitude. Fig. 5a shows the average of mean amplitude values across colors at the word-level. Labels for blue (M=81.05; SD=1.47) were significantly higher in amplitude than those for green (M=79.97; SD=1.54, p=.001), orange (M=78.57; SD=1.60, p<.001), purple (M=77.04; SD=1.99, p<.001), red (M=78.98; SD=1.93, p<.001), and yellow (M=78.87; SD=1.94, p<.001). Labels for green were significantly higher in amplitude than those for blue (p=.001), orange (p<.001), purple (p<.001), red (p=.003), and yellow (p=.003). Labels for orange were significantly higher in amplitude than those for purple (p<.001) but significantly lower in amplitude than those for blue (p<.001) and green (p<.001). Labels for purple were significantly lower in amplitude than those for blue (p<.001), green (p<.001), orange (p<.001), red (p<.001), and yellow (p<.001). Labels for red were significantly higher in amplitude than those for purple (p<.001), but significantly lower in amplitude than those for blue (p<.001) and green (p=.003). Labels for yellow were significantly higher in amplitude than those for purple (p<.001), but significantly lower in amplitude than those for blue (p<.001) and green (p=.003).

Duration. Fig. 6a shows the average of duration values across colors at the word-level. Labels for blue (M=.25; SD=.03) were significantly shorter in duration than those for green (M=.29; SD=.05, p<.001), orange (M=.39; SD=.05, p<.001), purple (M=.37; SD=.04, p<.001), and yellow (M=.37; SD=.03, p<.01), but significantly longer in duration than those for red (M=.23; SD=.03, p=.001). Labels for green were significantly shorter in duration than those for orange (p<.001), purple (p<.001), and yellow (p<.001), but significantly longer in duration than those for blue (p<.001) and red (p<.001). Labels for orange were significantly longer in duration than those for blue (p<.001), green (p<.001), purple (p=.02), red (p<.001), and yellow (p=.002). Labels for purple were significantly longer in duration than those for blue (p<.001), green (p<.001), and red (p<.001), but significantly shorter in duration than those for orange (p=.02).

Labels for red were significantly shorter in duration than those for blue (p=.001), green (p<.001), orange (p<.001), purple (p<.001), and yellow (p<.001). Labels for yellow were significantly longer in duration than those for blue (p<.001), green (p<.001), and red (p<.001), but significantly shorter in duration than those for orange (p=.002).

Differences in pitch across color labels suggested that acoustic properties may have changed as a function of color warmth. Color warmth in the current experiment was defined using the wavelength associated with each color, with 577nm as the borderline separating warm (red, orange, yellow) and cool (purple, blue, green) colors. Thus, colors associated with shorter wavelength (390nm- 577nm) were categorized as cool colors, whereas colors associated with longer wavelength (577nm – 780nm) were categorized as warm colors. To assess the extent to which labels differed as a function of color warmth, paired-samples *t*-tests were conducted to compare the mean F0, amplitude, and duration for warm and cool color labels. Results indicated a significant difference between cool (F_0, M=168.41, SD=38.95; Amplitude, M=79.35, SD=1.50; Duration, M=.31, SD=.04) and warm colors (F_0, M=163.13, SD=37.07; Amplitude, M=78.81, SD=1.66; Duration, M=.33, SD=.03) for mean F_0 ($t(19)$=4.50, p<.001), mean amplitude ($t(19)$=3.52, p=.002), and duration ($t(19)$= -4.83, p<.001). These results indicated that labels for cool colors were higher in pitch and amplitude but shorter in length compared to labels for warm colors.

Analyses on sentence-level utterances revealed similar results. Three one-way repeated-measures ANOVAs with brightness levels as within-subjects factor did not reveal any significant difference across the light, dark, and intermediate shades (F_0, F (2,38)=2.06, p=.14, *partial* η^2=.98; Amplitude, F (2,38)=.27, p=.76, *partial* η^2=.01; Duration, F (2,38)=.20, p=.82, *partial*

η^2=.01; see Fig 1b, 2b, 3b). However, mean F_0 for the lightest shades (*M*=175.82; *SD*=40.58) was higher than that for the darkest shades (*M*=174.93; *SD*=41.34; see Fig. 1b).

Three three-way repeated measures ANOVAs with brightness level, color, and block as within-subjects factors indicated a main effect of color for each of the three dependent measures (F_0, F (5,95)=3.19, p=.01, *partial* η^2=.14; Amplitude, F (5,95)=10.86, p<.001, *partial* η^2=.36; Duration, F (5,95)=51.20, p<.001, *partial* η^2=.73). A main effect of block was also found for duration (F (2,38)= 25.76, p<0.001, *partial* η^2=.58). However, no main effect of block for F_0 or amplitude was found (F_0, F (2,38)=2.03, p=.15, *partial* η^2=.10; Amplitude, F (2,38)=.16, p=.85, *partial* η^2=.01). No main effect of brightness level was found (F_0, F (2,38)=2.06, p=.14, *partial* η^2=.10; Amplitude, F (2,38)=.28, p=.76, *partial* η^2=.01; Duration, F (2,38)=.27, p=.77, *partial* η^2=.01), and no significant interactions were found. Pairwise comparisons of colors were conducted on all three dependent measures and were adjusted for multiple comparisons. These analyses are reported below.

Pitch. Fig. 4b shows the average of mean F_0 values across colors at the sentence-level. Labels for blue (*M*=175.00; *SD*=42.02) were significantly higher pitched than those for red (*M*=172.17; *SD*=40.93, p=.04). Labels for green (*M*=175.26; *SD*=41.25) were significantly higher pitched than those for red (p=.03). Labels for orange (*M*=175.69; *SD*=39.40) were not significantly different from any other colors in F_0. Labels for purple (*M*=177.87; *SD*=42.89) were significantly higher pitched than those for red (p=.002) and yellow (*M*=173.90; *SD*=40.46, p=.01). Labels for red were significantly lower pitched than those for blue (p=.04), green (p=.03), and purple (p=.002). Labels for yellow were significantly lower pitched than those for purple (p=.01).

Amplitude. Fig. 5b shows the average of mean amplitude values across colors at the sentence-levels. Labels for blue (*M*=76.53; *SD*=2.22) were significantly higher in amplitude than those for green (*M*=76.00; *SD*=2.13, $p<.001$), orange (*M*=75.85; *SD*=2.08, $p<.001$, purple (*M*=75.51; *SD*=2.44, $p<.001$), and red (*M*=75.74; *SD*=2.31, $p<.001$). Labels for green were significantly higher in amplitude than those for purple ($p=.01$), but significantly lower in amplitude than those for blue ($p<.001$). Labels in orange were significantly lower in amplitude than those for blue ($p<.001$) and yellow (*M*=76.39; *SD*=2.21, $p<.001$). Labels for purple were significantly lower in amplitude than those for blue ($p<.001$), green ($p=.01$), and yellow ($p<.001$). Labels for red were significantly lower in amplitude than those for blue ($p<.001$) and yellow ($p<.001$). Labels for yellow were significantly higher in amplitude than those for orange ($p<.001$), purple ($p<.001$), and red ($p<.001$).

Duration. Fig. 6b shows the average of duration values across colors at the sentence-level. Labels for blue (*M*=1.19; *SD*=.17) were significantly shorter in duration than those for green (*M*=1.23; *SD*=.17, $p<.001$), orange (*M*=1.32; *SD*=.19, $p<.001$), purple (*M*=1.26; *SD*=.17, $p<.001$), and yellow (*M*=1.25; *SD*=.15, $p<.001$). Labels for green were significantly longer in duration than those for blue ($p<.001$) and red (*M*=1.17; *SD*=.16, $p<.001$) but shorter in duration than those for orange ($p<.001$) and purple ($p<.001$). Labels for orange were significantly longer in duration than those for blue ($p<.001$), green ($p<.001$), purple ($p<.001$), red ($p<.001$), and yellow ($p<.001$). Labels for purple were significantly longer in duration than those for blue ($p<.001$), green ($p<.001$), and red ($p<.001$), but shorter in duration than those for orange ($p<.001$). Labels for red were significantly shorter in duration than those for green ($p<.001$), orange ($p<.001$), purple ($p<.001$), and yellow ($p<.001$). Labels for yellow were significantly longer in duration than those for blue ($p<.001$) and red ($p<.001$), but shorter in duration than those for orange ($p<.001$).

As with the single-word level analyses, there was a significant difference between cool (F_0, M=176.04, SD=41.90; Amplitude, M=76.02, SD=2.23; Duration, M=1.23, SD=.17) and warm (F_0, M=173.92, SD=40.05; Amplitude, M=75.99, SD=2.16; Duration, M=1.25, SD=.16) colors for mean F_0 ($t(19)=2.65$, $p=.02$) and duration ($t(19)= -2.89$, $p=.01$). The difference between warm and cool colors was not significant for mean amplitude at the sentence level ($t(19)=.19$, $p=.85$), but the trend was consistent with the results at word-level, in that cool colors had higher mean amplitude than warm colors.

Taken together, the results for both word- and sentence-level analyses suggest that participants did not reliably vary their prosodic cues in labels for different brightness levels. However, participants did reliably vary their prosodic cues across labels for different colors. In particular, labels for cool colors were produced with higher pitch and louder amplitude compared to those for warm colors. The main effect of color for duration in both word- and sentence-level analyses may have reflected the number of syllables included in the color labels (e.g., "blue" has one syllable but "orange" has two syllables).

One possible explanation for this pattern of results may be that the pronunciations of the cool color names themselves tend to lead to higher pitch responses. Previous studies have shown that certain vowels, typically high front vowels, have higher F_0 than other vowels (e.g., Hillenbrand et al., 1995). Hillenbrand et al. (1995) found that different vowels are associated with different sets of acoustic profiles. For example, /i/, /I/, /U/, and /u/ in American English are associated with higher F_0 across female, male, and child speakers (Hillenbrand et al., 1995). Both the color names "green" (/grin/) and "blue" (/blu/) have relatively high F_0 vowels and were associated with higher pitch utterances in the present experiment. However the word "orange" (/ɔrɪndʒ/) also has a high F_0 vowel but was associated with lower pitch utterances in the current experiment. Further, the word "purple" (/pərpəl/), which had one of the highest mean F_0 across participants in our study, does not contain any high F_0 vowel sounds. Thus, although the acoustic features of the vowels in the color labels may have contributed to speakers' variations in pitch and amplitude, the prosodic cues in speakers' color labels cannot be fully attributed to the typical acoustic features of specific vowels, as high F_0 vowels did not necessarily correspond to high pitch utterances.

Experiment 2

Although the productions of color names in Experiment 1 did not appear to be based solely on the typical acoustic properties of vowel sounds, the presence of the color names may still have affected how speakers represented color brightness in their utterances. The objective of Experiment 2 was to examine the extent to which the color labels influenced the production of prosodic cues in speakers' utterances. Further, Experiment 2 examined whether prosodic cues to brightness and color would emerge when linguistic content is ambiguous or underdetermined. To control for the effects that color names may have on participants' responses, we replaced the color labels used in Experiment 1 with a set of six bi-syllabic novel words (*blicket, daxen, foppick, tillen, riffel, seebow*). Language users have no previous experience associating these novel words with particular referents as they do for the actual color names. Participants participated in a production task similar to the one in Experiment 1, where they were asked to produce verbal labels for different brightness levels of various colors. As for Experiment 1, we predicted that speakers would produce higher pitch utterances for the lightest shade of a particular color than the darkest shade, and that the prosodic realization of verbal labels would vary significantly across colors. Further, we hypothesized that the acoustic properties of cool colors would differ significantly from the acoustic properties of warm colors.

Methods

Participants. Thirty-two Emory University undergraduate students participated in the experiment. Analyses were conducted with data from 21 participants (11 participants were excluded due to familiarity with a foreign language before the age seven). All participants were monolingual American English speakers with no reported history of speech or hearing disorders. Participants were recruited from the Emory University Introductory Psychology subject pool and received course credit for participation.

Materials. The same set of color spectrums from Experiment 1 were used. However, the color labels provided above the arrows were changed to the novel word labels, *blicket, daxen, foppick, tillen, riffel, seebow*.

Procedure. Participants completed the same production task as Experiment 1, in which they were asked to provide spoken labels for the lightest, darkest, and an intermediate shade of each of the six colors. To rotate the pairing between the novel word labels and their associated colors, six experimental conditions were created. In each condition, the novel words were associated with a different color (e.g., for Condition 1, “blicket” was associated with the color blue; whereas for Condition 2, “blicket” was associated with the color yellow). All other aspects of the procedure remained the same as those described in Experiment 1.

Acoustic Measures. For each sentence and word utterance, acoustic measures were obtained to determine if participants’ productions varied as a function of brightness and color. Mean F_0, mean amplitude and duration were measured using PRAAT.

Results and Discussion

Separate acoustic analyses were conducted on both sentence- and word-level utterances. To assess the overall performance across brightness levels for F_0, amplitude and duration, three one-way repeated-measures ANOVAs were conducted with brightness level as a within-subjects factor on word-level utterances. Overall, the analyses did not reveal a significant difference among brightness levels across any of the three dependent measures (F_0, F (2,40)=2.34, p=.11, *partial* η^2=.11; Amplitude, F (2,40)=0.51, p=.61, *partial* η^2=.03; Duration, F (2,40)=2.50, p=.09, *partial* η^2=.11; see Fig. 7a, 8a, 9a). However, mean F_0 for the lightest shades (M=170.69; SD=38.17) was higher than that for the darkest shades (M=166.10; SD=36.50; see Fig. 7a).

Three three-way repeated-measures ANOVAs with brightness levels, color, and block as within-subjects factors did not yield any significant main effects or interactions for F_0 and amplitude. A main effect of block for duration was found (F (2,40)=7.07, p=.002, *partial* η^2=.26). Pairwise comparisons indicated that utterances in block 1 (M=.48; SD=.06) were significantly longer in duration than those in block 2 (M=.45; SD=.05, p=.001) and block 3 (M=.44; SD=.05, p=.01), suggesting that the lengths of utterances decreased across blocks. No main effect of brightness level or color was found for duration (Brightness level, F (2,40)=2.50, p=.09, *partial* η^2=.11; Color, F (5,100)=.68, p=.64, *partial* η^2=.03).

Although prosodic cues did not appear to vary as a function of color, given the results from Experiment 1, exploratory analyses were conducted on colors to examine the differences across colors that were found in Experiment 1. Pairwise comparisons across colors indicated that labels for purple (M=163.74; SD=36.58) were significantly lower pitched than those for red (M=170.82; SD=37.03, p=.03; see Fig. 10a). Labels for blue (M=.44; SD=.08) were significantly shorter in duration than those for yellow (M=.48; SD=.08, p=.04; see Fig. 12a). No other significant differences were found across colors in all three dependent measures (see Fig. 10a, 11a, 12a). To examine the extent to which labels differed as a function of color warmth, paired-samples *t*-tests were conducted. Cool colors (F_0, M=166.20, SD=36.82; Amplitude, M=76.17, SD=2.20; Duration, M=.45, SD=.05) did not differ significantly from warm colors (F_0, M=168.50, SD=37.11; Amplitude, M=76.11, SD=2.21; Duration, M=.46, SD=.06) for all three dependent measures (F_0, $t(20)$= -1.09, p=.29; Amplitude, $t(20)$=.16, p=.87; Duration, $t(20)$= -.43, p=.67).

Analyses on sentence-level utterances revealed similar results. Three one-way repeated-measures ANOVAs with brightness level as within-subjects factor did not reveal any significant difference across brightness levels (F_0, F (2,40)=2.29, p=.11, *partial* η^2=.10; Amplitude, F (2,40)=1.85, p=.17, *partial* η^2=.09; Duration, F (2,40)=1.03, p=.37, *partial* η^2=.05; see fig. 7b, 8b, 9b). However, mean F_0 associated with the lightest shades (M=176.52; SD=41.51) was higher than that for the darkest shades (M=173.28; SD=40.34; see Fig. 7b).

Three three-way repeated measures ANOVAs with brightness level, color, and block as factors did not yield any significant main effect or interactions for F_0 and amplitude. A main effect of block was found for duration (F (2,40)=12.17, p<.001, *partial* η^2=.38). Pairwise comparisons indicated that utterances in block 1 (M=1.54; SD=.25) were significantly longer in duration than those in block 2 (M=1.46; SD=.22, p=.001) and block 3 (M=1.39; SD=.19, p=.001), and those in block 2 were significantly longer in duration than those in block 3 (p=.02), suggesting that the lengths of utterances decreased over blocks. Interestingly, an independent-samples *t*-test revealed that the duration of utterances in the current experiment (M=1.47; SD=.21) was significantly longer than that in Experiment 1 (M=1.24; SD=.17; $t(39)$= -3.88, p<.001), suggesting that the current task, given the use of novel words, may have been more difficult than the task in Experiment 1. No main effect of brightness level or color was found for duration (Brightness level, F (2,40)=1.03, p=.37, *partial* η^2=.05; Color, F (5,100)=.20, p=.96, *partial* η^2=.01).

Exploratory pairwise comparisons across colors indicated that labels for purple (M=171.72; SD=40.21) were significantly lower pitched than those for green (M=175.064; SD=41.38, p=.02), orange (M=175.52; SD=40.03, p=.03), red (M=175.39; SD=40.22, p=.02), and yellow (M=174.46; SD=41.27, p=.04; see fig. 10b). No other significant differences were found for colors across all three dependent measures (see Fig. 10b, 11b, 12b). Paired-samples t-tests yielded no significant difference between warm (F_0, M=175.13, SD=40.22; Amplitude, M=74.77, SD=2.69; Duration, M=1.47, SD=.20) and cool (F_0, M=173.66, SD=40.98; Amplitude, M=74.83, SD=2.71; Duration, M=1.46, SD=.21) colors across all three dependent measures (F_0, t(20)= -1.29, p=.21; Amplitude, t(20)=.35, p=.73; Duration, t(20)= -.58, p=.57).

The results for both word- and sentence-level analyses suggest that participants did not reliably vary their prosodic cues across brightness levels. However, participants did overall produce higher F_0 for lighter shades than darker shades. In the current experiment, participants did not vary their prosodic cues for labels across colors. However, exploratory analyses indicated that labels for purple were lower in pitch than other colors at both the word- and sentence-levels. This is the reverse of the pattern observed in Experiment 1, where labels for cool colors were higher in F_0 than those for warm colors. One possible explanation would be that purple is in general a perceptually darker color. Thus the presentation of a purple spectrum elicited lower pitched utterances. In fact, Marks et al. (1987) found that participants judged perceptually darker colors (e.g., blue and brown) to be lower pitched than perceptually brighter colors (e.g., yellow). Further, participants did not vary their acoustic cues across warm and cool colors. Unlike in Experiment 1, the mean F_0 for warm colors was higher, albeit non-significantly higher, than that for the cool colors at both word- and sentence-levels. This observation may be consistent with Marks et al.'s study (1987), in which colors such as red and yellow were judged by participants to be higher pitched than colors such as green and blue.

General Discussion

The purpose of the current investigation was to examine how language users incorporate prosodic cues to convey meaning-based information in their speech. In particular, we sought to investigate the extent to which language users are sensitive to cross-modal correspondences between visual and auditory domains and whether their utterances would reflect underlying mappings between acoustic properties and color brightness. In Experiment 1, participants did not reliably alter their prosodic cues in verbal labels across levels of brightness. However, speakers varied their prosodic cues in labels across colors. Specifically, they produced acoustic profiles with higher pitch and louder amplitude for cool colors than for warm colors.

In Experiment 2, we investigated whether the observed cross-modal mappings in Experiment 1 would be consistent when color name labels were arbitrarily related to colors. When participants labeled colors with novel words, they did not reliably vary their prosodic cues in labels across brightness levels. However, the difference between the mean F_0 level associated with brighter shades and that associated with darker shades was greater than that found in Experiment 1. In Experiment 2, participants did not vary their prosodic cues for labels across colors, and acoustic profiles for warm and cool colors were not significantly different from each other. Exploratory analyses examining differences across colors and warm and cool categorizations revealed an opposite trend to that exhibited by speakers in Experiment 1. In the second experiment, labels for warm colors were in general higher in pitch than those for cool colors. Further, labels for purple were found to be significantly lower in pitch than those for green, orange, red and yellow at the sentence-level.

One possible explanation for this discrepancy between the two experiments may be due to the fact that there is rich conceptual information attached to the actual color names. The fact that the color names have been long associated with their corresponding visual colors may elicit semantic information or concepts that extend beyond the perceptual properties of the visual colors. For example, previous research in color perception has shown that warm colors such as red, orange, and yellow are often associated with images of fire, while cool colors, in particular blue and green, are associated with images of water and air in art and fashion (Kim, 2006). Thus, the prosodic cues elicited for different colors in Experiment 1 may not necessarily be attributed to the perceptual properties of the colors but rather to the metaphorical associations strongly connected with these specific colors. These extensive metaphorical associations to the colors labels may have elicited utterances that extend beyond lower-level perceptual mappings. For example, the association between the color blue and the image of air may elicit higher pitched responses due to this conceptual image of an element that is vertically up or light in weight. In alignment with this assumption, although the visual and auditory mappings in Experiment 1 were inconsistent with Marks et al.'s study (1987), when color labels were removed in Experiment 2, utterances seemed to show a trend that was similar to that observed in Marks et al.'s study (e.g., labels for red, yellow, and orange were higher pitched than those for purple). This suggests that the extensive semantic knowledge associated with color labels may have played a role in how speakers were producing prosodic cues in speech.

Another possibility for the discrepancy between the two experiments is the fact that the task in Experiment 2 may be more complex in nature than that in Experiment 1. For example, participants may have been overly focused on the pronunciations of the novel words, which in turn interfered with their productions of prosodic cues in relationship to changes across levels of brightness and colors. In fact, the duration of the sentence utterances was significantly longer in Experiment 2 than that in Experiment 1. Further, in Experiment 2, the duration of utterances was significantly longer in block 1 than it was in block 2 and 3, suggesting that initially participants may have viewed the task as more difficult. This may have impacted how participants were responding to the color spectrums.

Aside from the discrepancies between the two experiments, findings in brightness levels were inconsistent with previous empirical findings in that both experiments indicated that participants did not reliably alter their prosodic cues in labels across different levels of brightness. Thus, in comparison to darker shades, lighter shades did not reliably map onto higher pitch utterances. However, across both experiments the mean F_0 for lighter shades was higher than that for darker shades (see Fig. 1a and 7a). Further, when labels were arbitrarily assigned to colors in Experiment 2, the difference of mean F_0 between the darkest and lightest shades increased in the predicted direction (see Fig. 7a). This suggests that the actual color names may have also interfered with how speakers were responding to brightness levels within colors. Thus, potential semantic or metaphorical associations attached to the color names may have elicited responses that correspond to properties that are beyond perceptual brightness.

This inconsistency with previous findings may also be due to the fact that the tasks employed in the current study were too explicit in nature. For example, because participants were directly asked to describe the shade of the color using only the provided labels, the explicit nature of the task might have somehow masked their spontaneous responses corresponding to levels of brightness. In contrast, Shintel et al. (2006) used more implicit tasks, in which the focus of the tasks was to describe the direction of movement of a dot (e.g., left or right) rather than the dimension of interest (e.g., speed). Thus, the implicit nature of their tasks may have distracted the participants from the real purpose of the study and allowed them to spontaneously exhibit prosodic correlates to visuo-spatial information.

Although findings from the present investigation did not support a sound-to-meaning mapping for prosodic correlates to brightness perceptions, they nonetheless suggest a non-arbitrary cross-modal correspondence between the visual and auditory domains in spoken language. Further research is needed to determine the specificities of such underlying mappings between the two domains. For example, it would be interesting to examine the extent to which these mappings are consistent across contexts. Taken together, current findings imply a complex relationship between the role of prosody and meaning in human communication. Different from traditional views that assume prosody to mainly assist in the comprehension of linguistic structures or reflect nonlinguistic features such as speaker intentionality or emotional state, the current findings suggest that prosody may directly carry visual-referential information in spoken language in that speakers' utterances may reflect their sensitivity to different levels of visual and auditory mappings.

References

Bachorowski, J. A. (1999). Vocal expression and perception of emotion. *Current Directions in Psychological Science, 8*, 53–57. doi:10.1111/1467-8721.00013

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577–609. doi:10.1017/S0140525X99002149

Boersma, Paul & Weenink, David (2012). Praat: doing phonetics by computer [Computer program]. Version 5.3.43, retrieved 27 March 2013 from http://www.praat.org/

Clifton, C. r., Carlson, K., & Frazier, L. (2006). Tracking the what and why of speakers' choices: Prosodic boundaries and the length of constituents. *Psychonomic Bulletin & Review, 13*(5), 854-861. doi:10.3758/BF03194009

Cutler, A. (1997). Prosody and the structure of the message. In Y. Sagisaka, N. Campbell & N. Higuschi (Eds.), *Computing prosody: Computational models for processing spontaneous speech* (pp. 63–66). New York: Springer Verlag.

de Saussure, F. (1966[1915]). *Course in general linguistics.* W. Baskin (trans.). C. Ballly & A. Sechehaye (Eds). New York: McGraw-Hill.

Herold, D. S., Nygaard, L. C., & Namy, L. L. (2012). Say it like you mean it: Mothers' use of prosody to convey word meaning. *Language And Speech, 55*(3), 423-436. doi:10.1177/0023830911422212

Hockett, C. F. (1960). The origin of speech. *Scientific American, 203*, 88–96.

Jusczyk, P. W., Hirsh Pasek, K., Nelson, D. G. K., Kennedy, L. J., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology, 24*, 252–293. doi:10.1016/0010-0285(92)90009-Q

Kim, Young-In. (2006). Color and symbolic meaning of elements in nature. *Color Research and*

Application, 31(4), 341-349. doi:10.1002/col.20232

Marks, L. E. (1982). Bright sneezes and dark coughs, loud sunlight and soft moonlight. *Journal of Experimental Psychology: Human Perception and Performance, S,* 177-193. doi:10.1037/0096-1523.8.2.177

Marks, L. E. (1987). On cross-modal similarity: Auditory–visual interactions in speeded discrimination. *Journal Of Experimental Psychology: Human Perception And Performance, 13*(3), 384-394. doi:10.1037/0096-1523.13.3.384

Marks, L. E., Hammeal, R. J., & Bornstein, M. H. (1987). Perceiving similarity and comprehending metaphor. *Monographs Of The Society For Research In Child Development, 52*(1), 1-92. doi:10.2307/1166084

Nygaard, L. C., Herold, D. S., & Namy, L. L. (2009). The semantics of prosody: Acoustic and perceptual evidence of prosodic correlates to word meaning. *Cognitive Science, 33,* 127–146. doi:10.1111/j.1551-6709.2008.01007.x

Nygaard, L. C., & Lunders, E. R. (2002). Resolution of lexical ambiguity by emotional tone of voice. *Memory & Cognition, 30*, 583–593. doi:10.3758/BF03194959

Ofuka, E., McKeown, J. D., Waterman, M. G., & Roach, P. J. (2000). Prosodic cues for rated politeness in Japanese speech. *Speech Communication, 32*, 199–217. doi:10.1016/S0167-6393(00)00009-1

Scherer, K. (1994). Vocal affect expression: A review and a model for future research. *Psychological Bulletin, 92*, 143–165. doi:10.1037/0033-2909.99.2.143

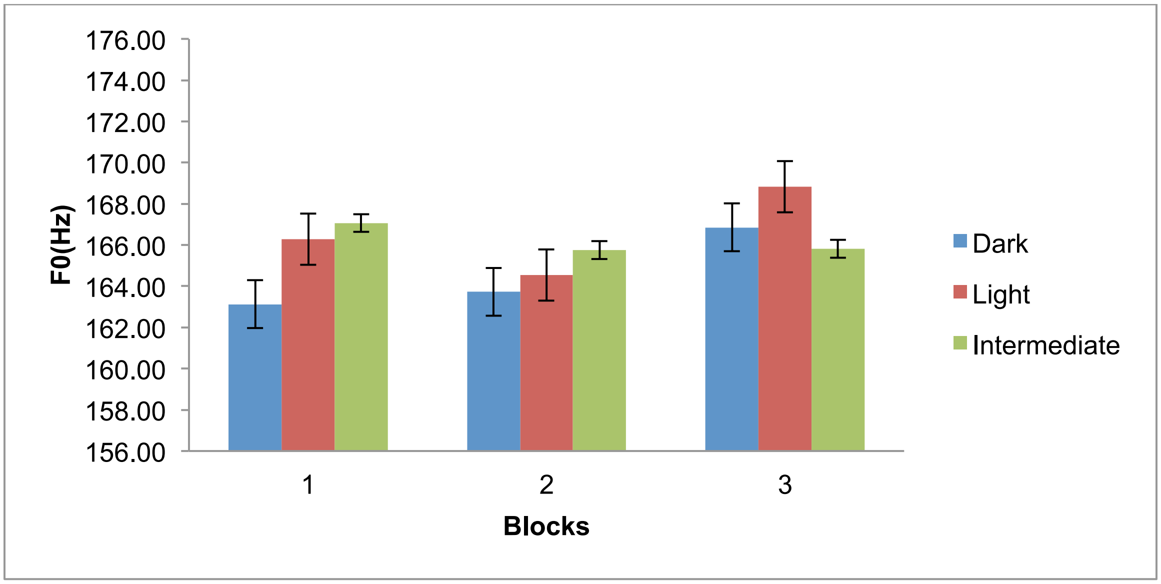
Shintel, H., & Nusbaum, H. C. (2007). The sound of motion in spoken language: Visual information conveyed by acoustic properties of speech. *Cognition, 105,* 681–690.

Shintel, H., Nusbaum, H. C., & Okrent, A. (2006). Analog acoustic expression in speech communication. *Journal of Memory and Language, 55*, 167–177.

Spence, M. J., & Moore, D. (2003). Categorization of infant-directed speech: Development from 4 to 6 months. *Developmental Psychobiology, 42*, 97–109.

Wurm, L. H., Vakoch, D. A., Strasser, M. R., Calin-Jageman, R., & Ross, S. E. (2001). Speech perception and vocal expression of emotion. *Cognition and Emotion, 15*, 831–852.

a)



b)

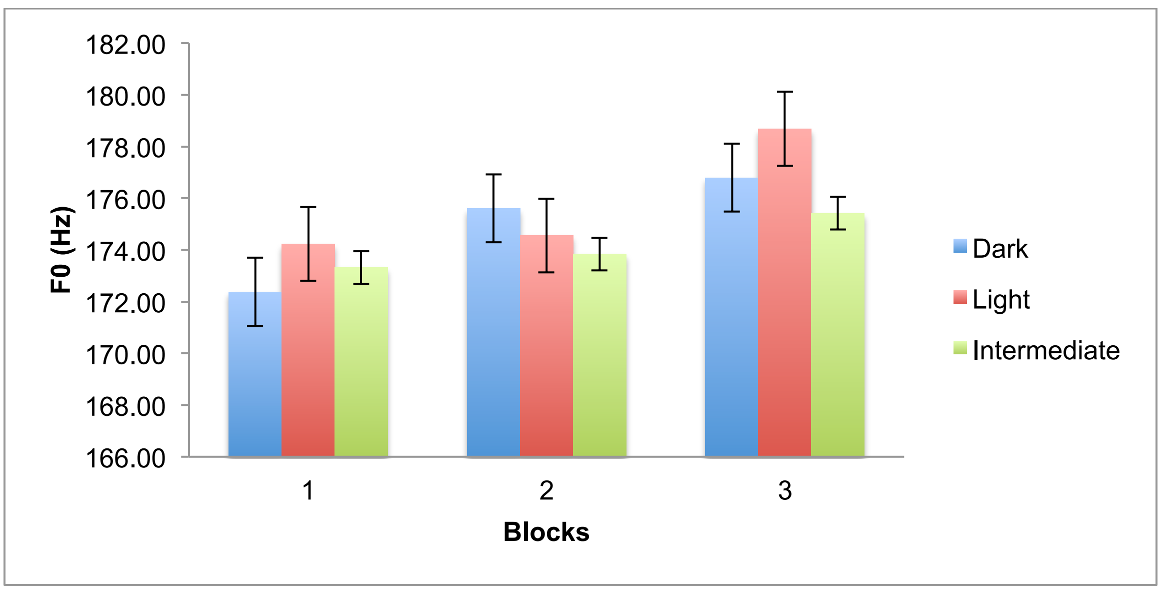
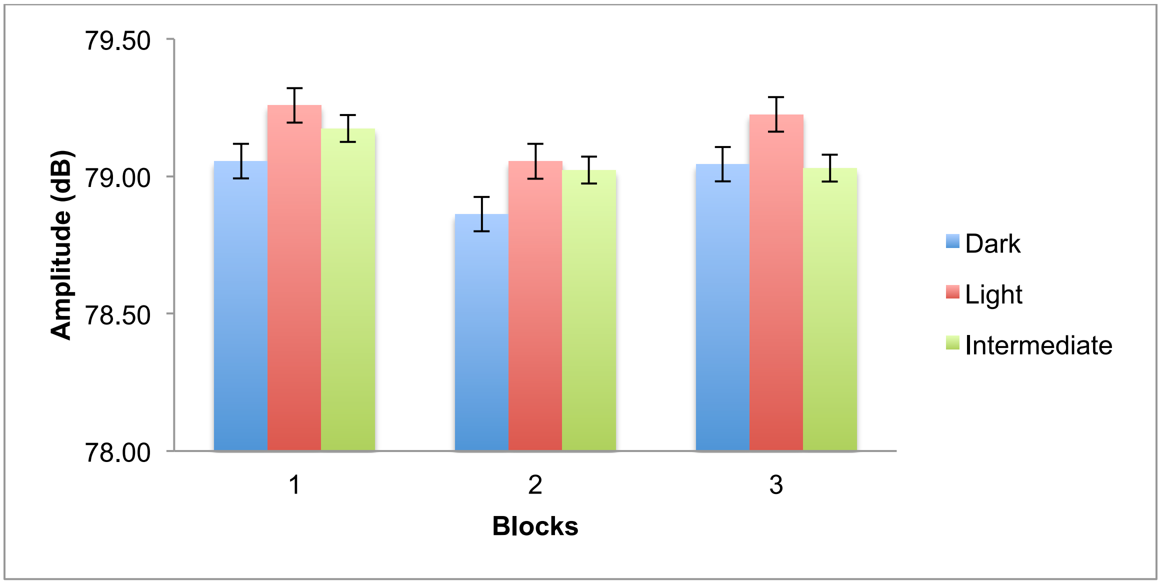


Figure 1. a) Mean F0 across brightness levels by blocks at word-level in Experiment 1. b) Mean F0 across brightness levels by blocks at sentence-level in Experiment 1. Error bars represent one standard error of the mean.

a)



b)

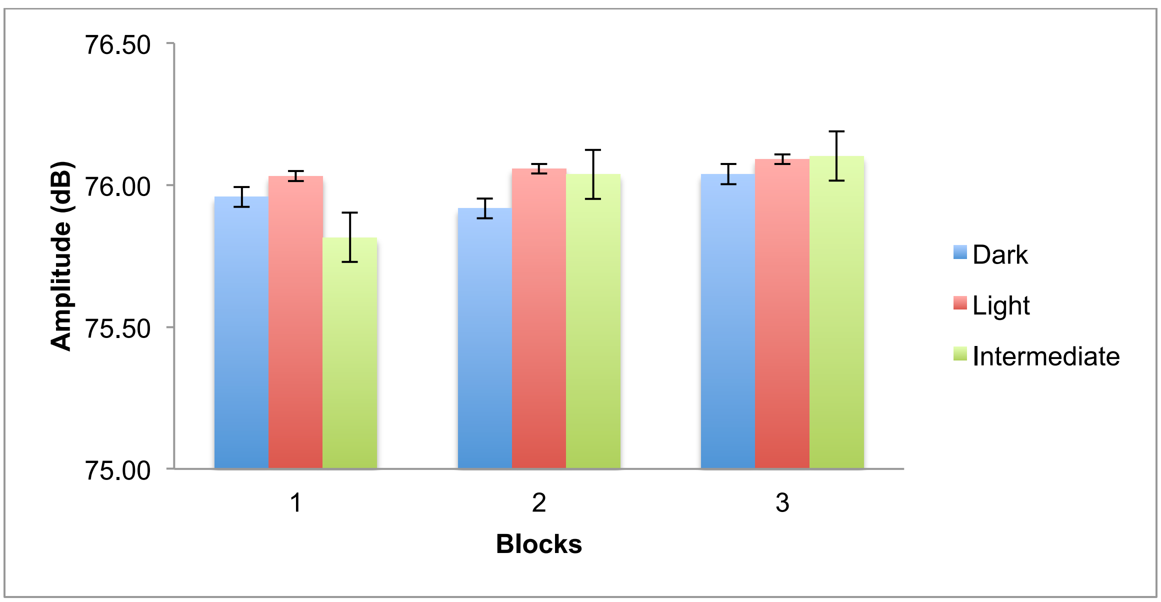
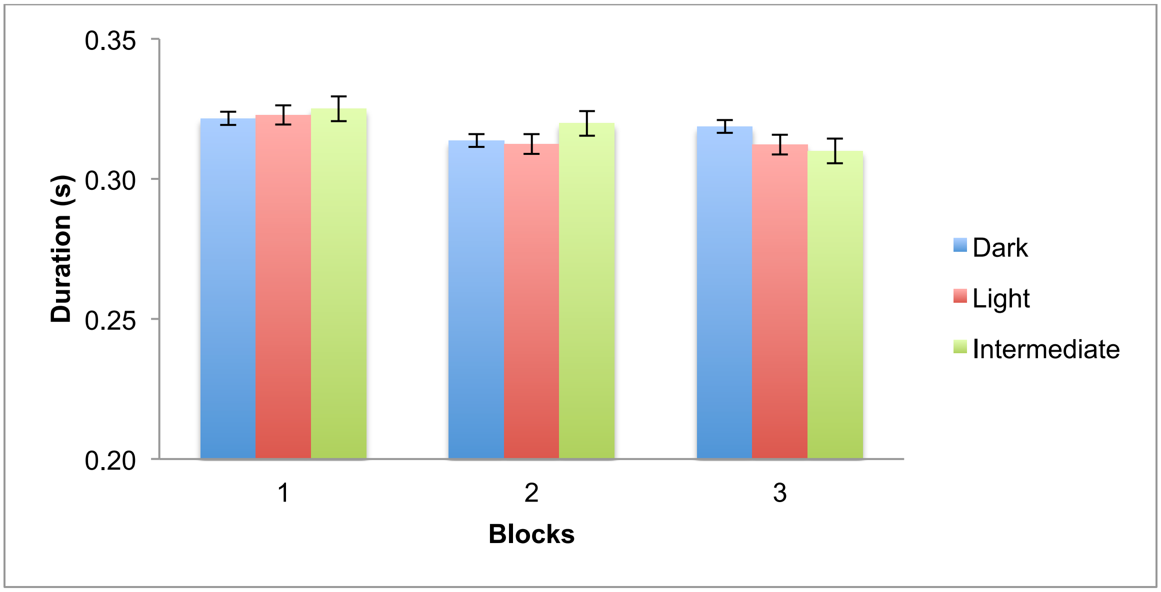


Figure 2. a) Mean amplitude across brightness levels by blocks at word-level in Experiment 1. b) Mean amplitude across brightness levels by blocks at sentence-level in Experiment 1. Error bars represent one standard error of the mean.

a)



b)

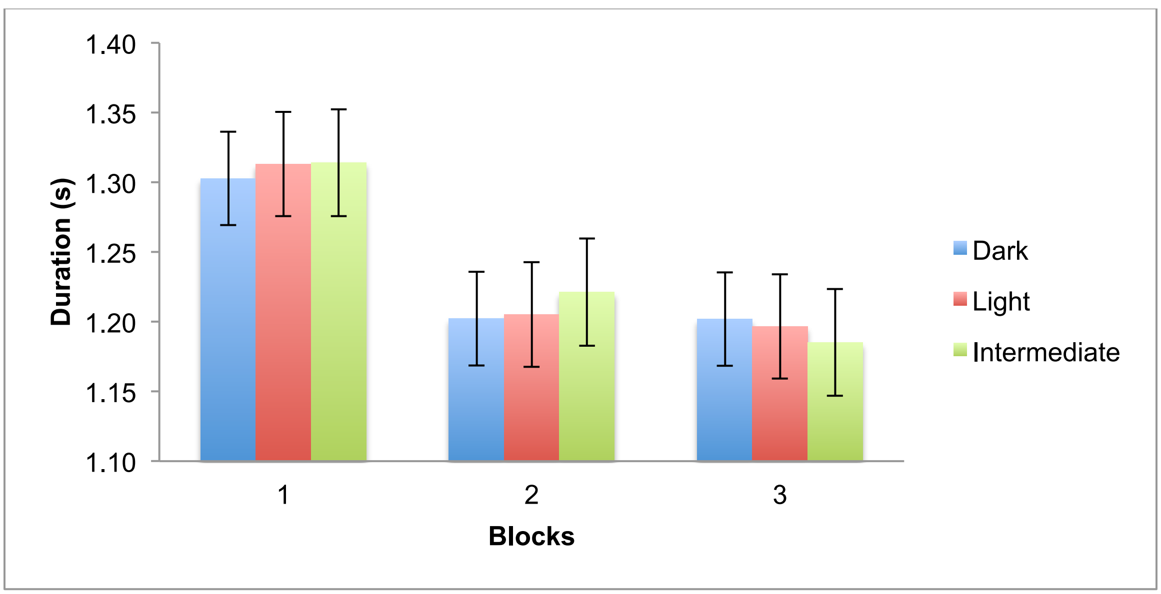
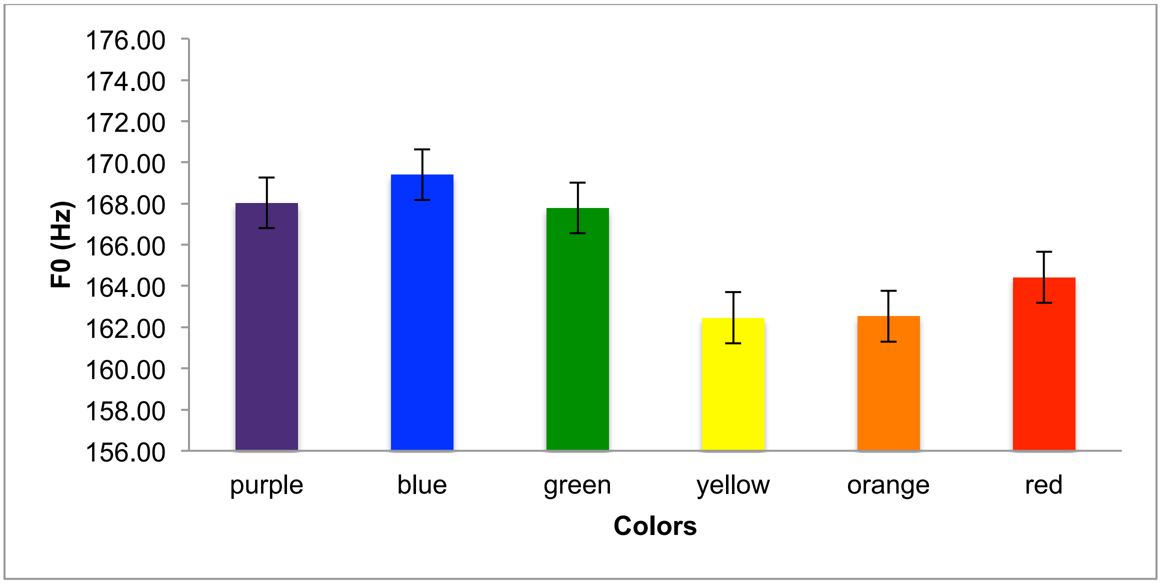


Figure 3. a) Average duration across brightness levels by blocks at word-level in Experiment 1. b) Average duration across brightness levels by blocks at sentence-level in Experiment 1. Error bars represent one standard error of the mean.

a)



b)

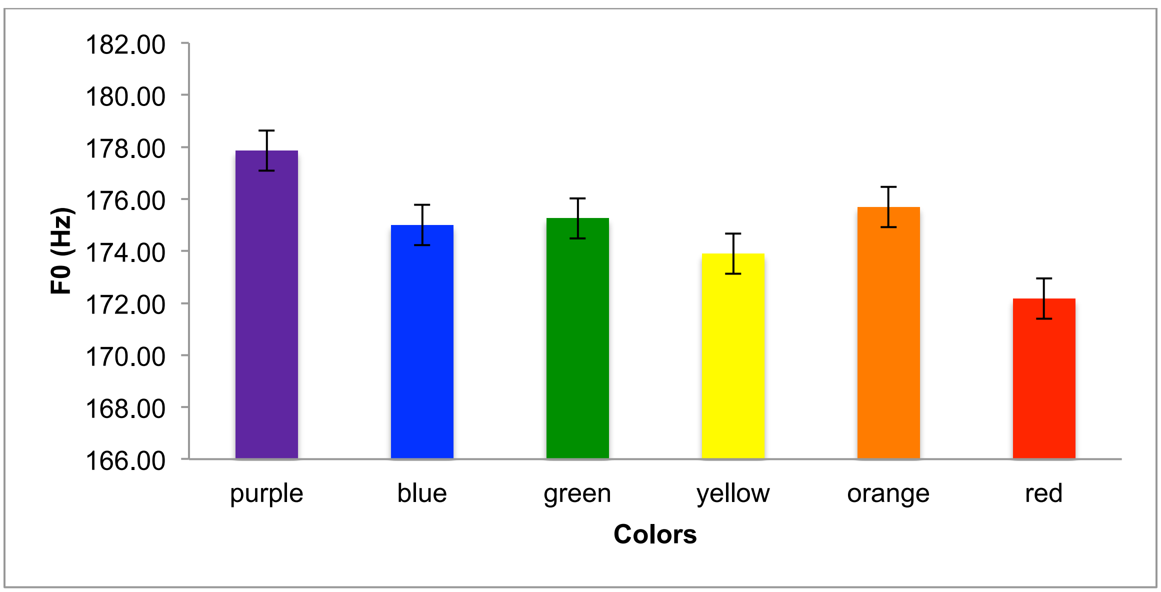
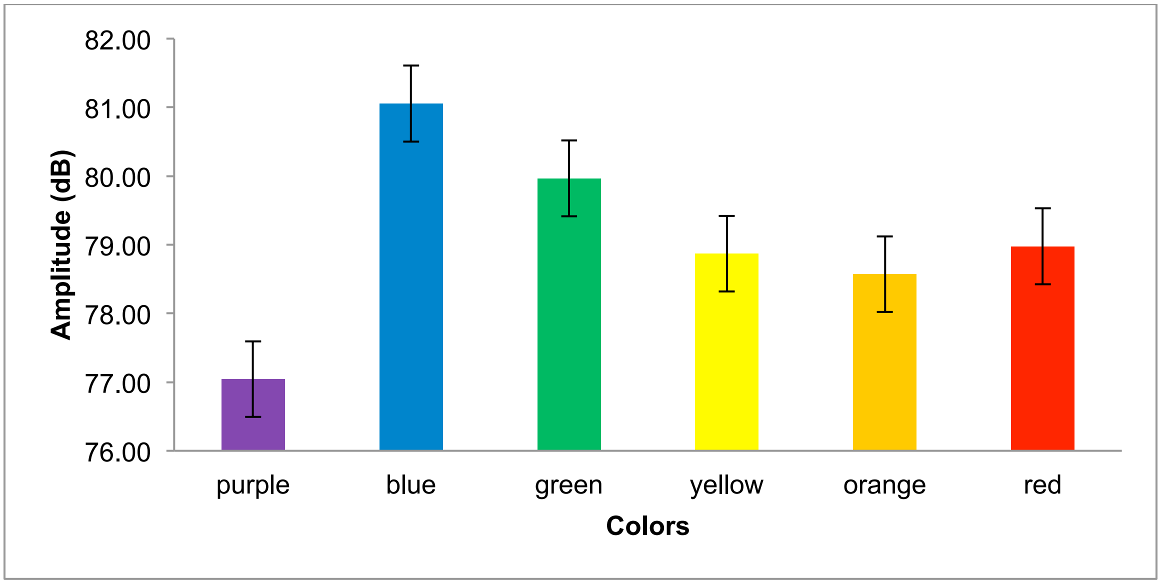


Figure 4. a) Mean F0 across colors at word-level in Experiment 1. b) Mean F0 across colors at sentence-level in Experiment 1. Error bars represent one standard error of the mean.

a)



b)

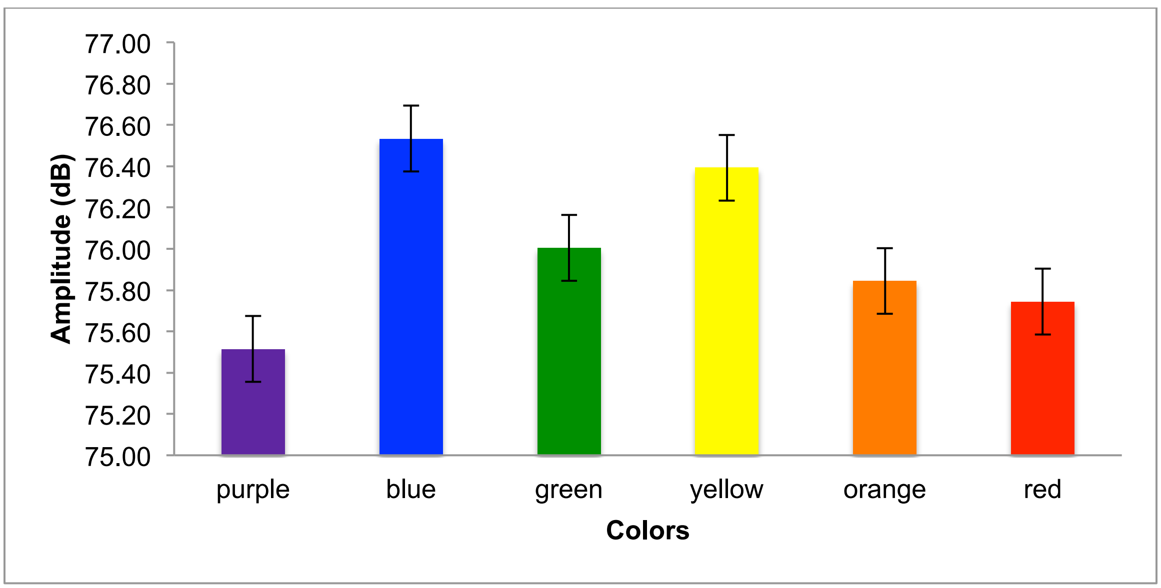
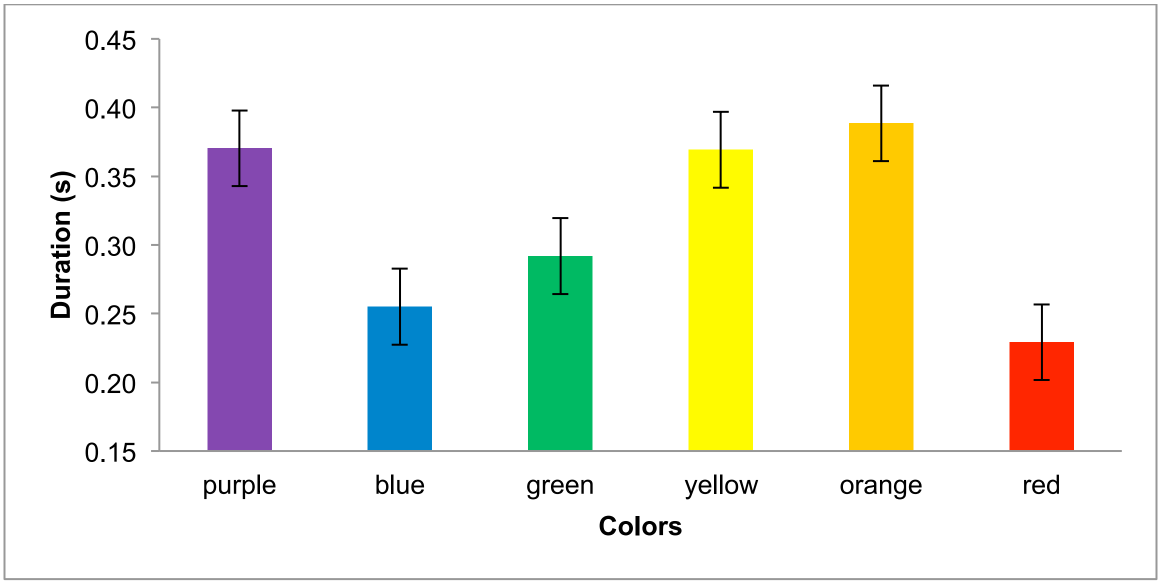


Figure 5. a) Mean amplitude across colors at word-level in Experiment 1.
b) Mean amplitude across brightness levels by blocks at sentence-level in Experiment 1. Error bars represent one standard error of the mean.

a)



b)

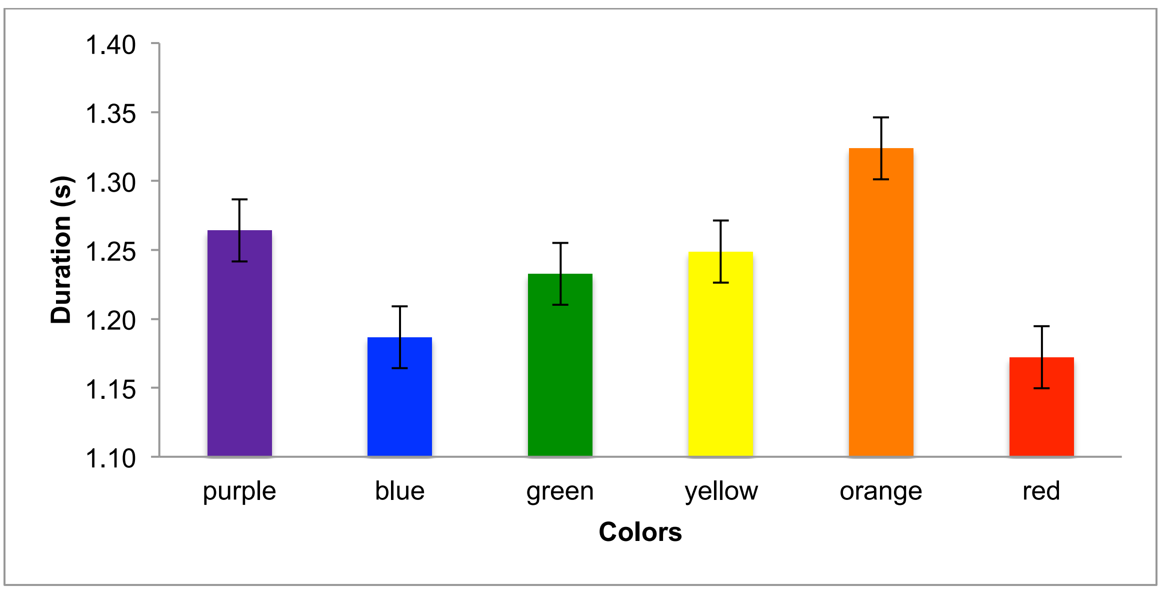
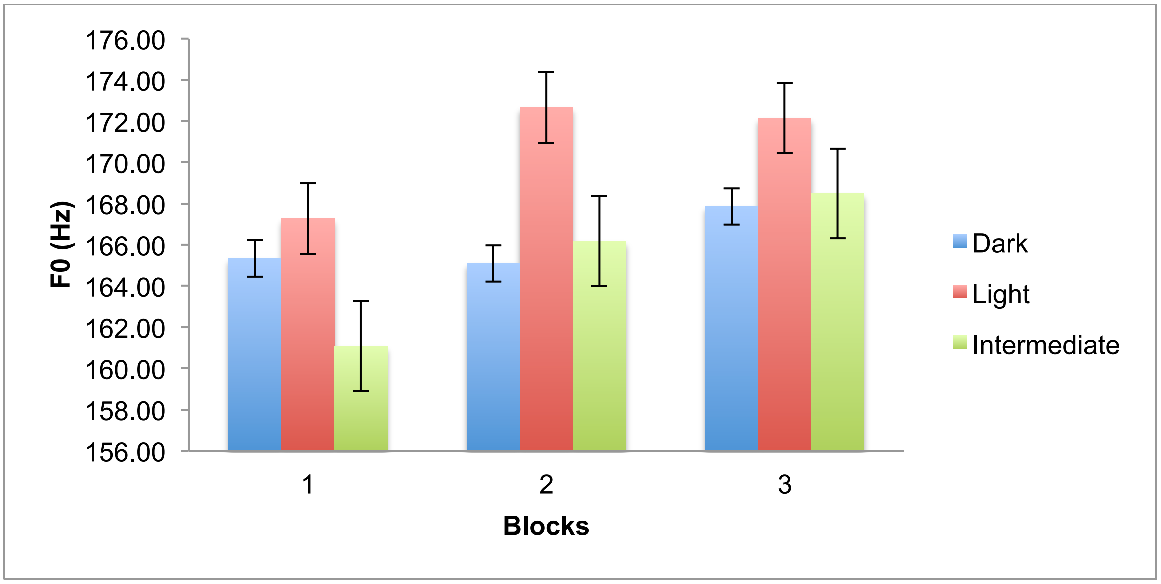


Figure 6. a) Average duration across colors at word-level in Experiment 1.
b) Average duration across colors at sentence-level in Experiment 1. Error bars represent one standard error of the mean.

a)



b)

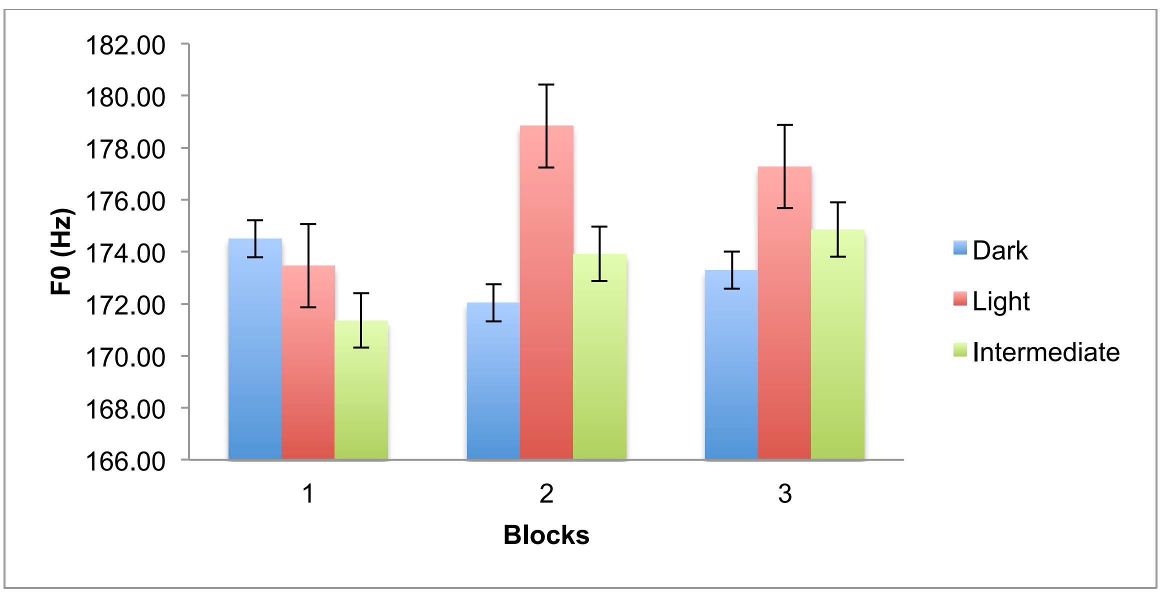
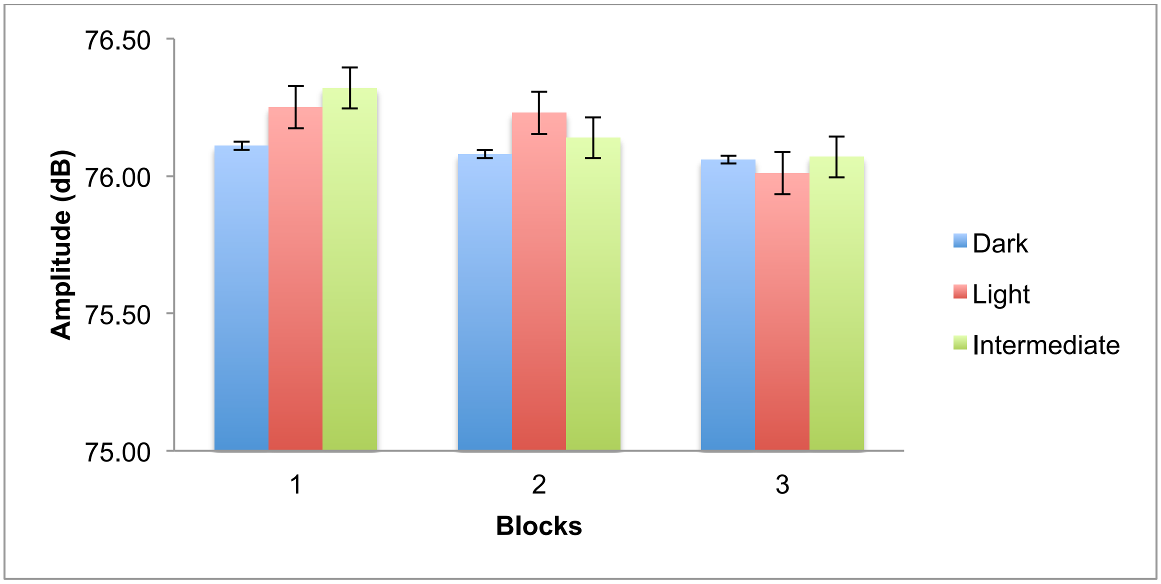


Figure 7. a) Mean F0 across brightness levels by blocks at word-level in Experiment 2. b) Mean F0 across brightness levels by blocks at sentence-level in Experiment 2. Error bars represent one standard error of the mean.

a)



b)

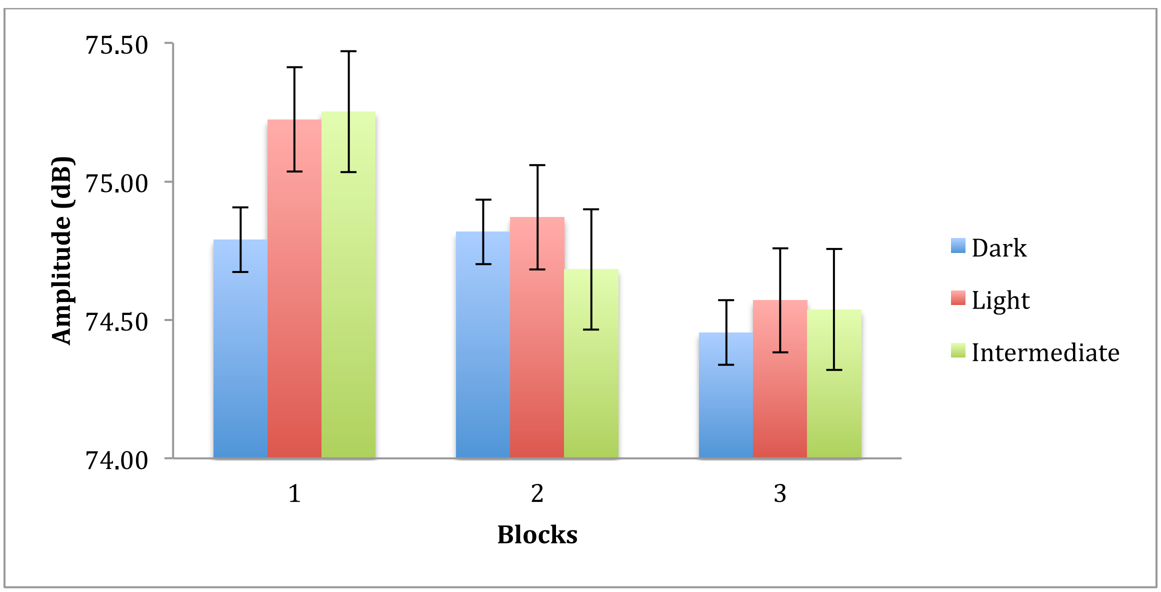
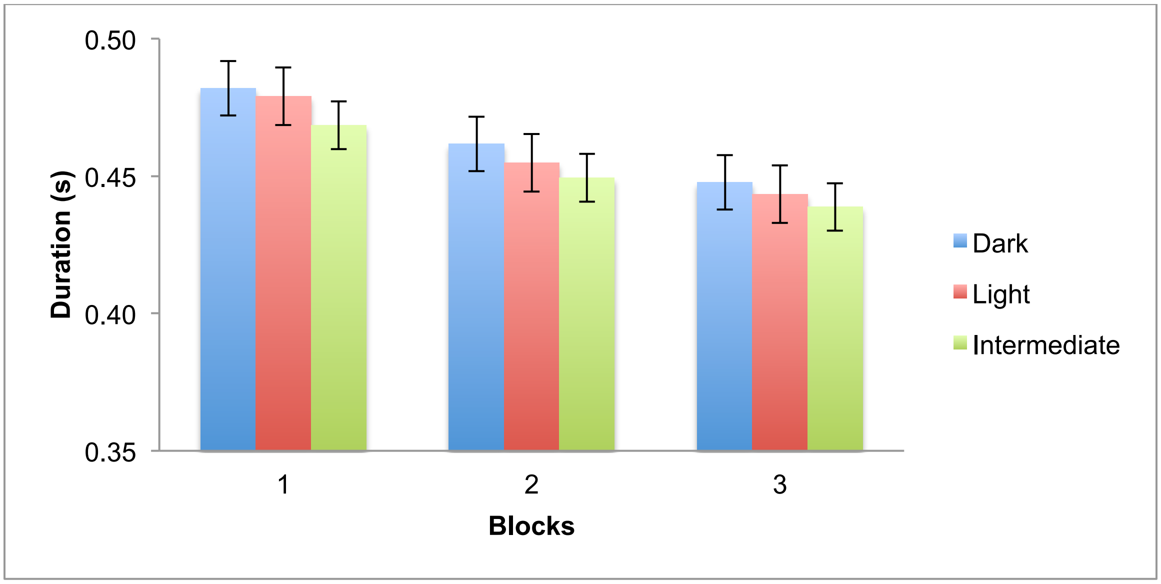


Figure 8. a) Mean amplitude across brightness levels by blocks at word-level in Experiment 2. b) Mean amplitude across brightness levels by blocks at sentence-level in Experiment 2. Error bars represent one standard error of the mean.

a)



b)

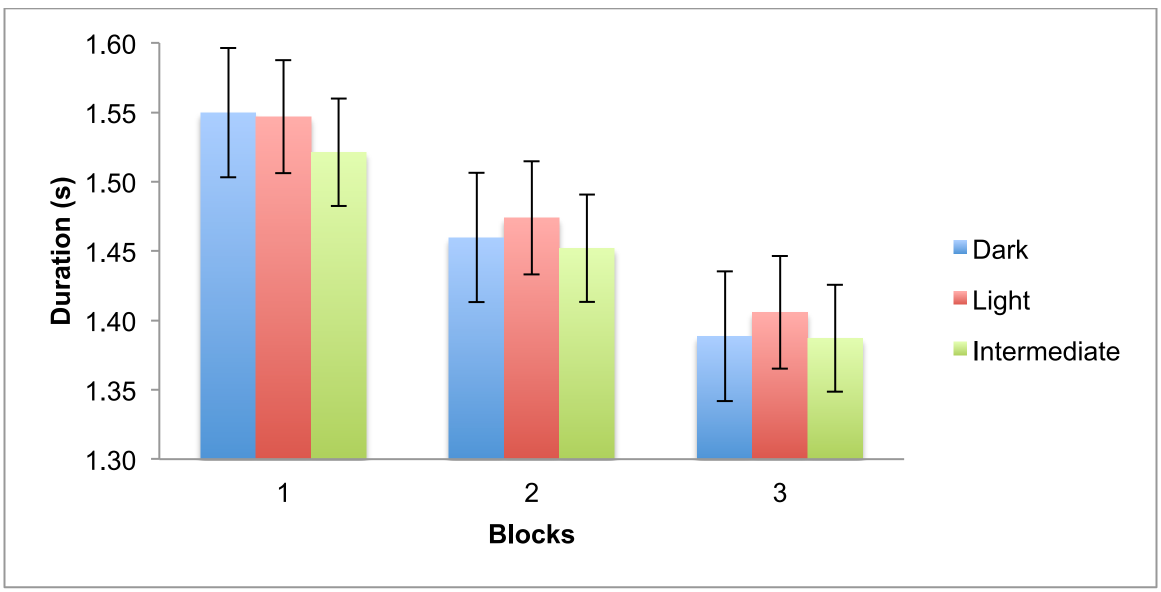
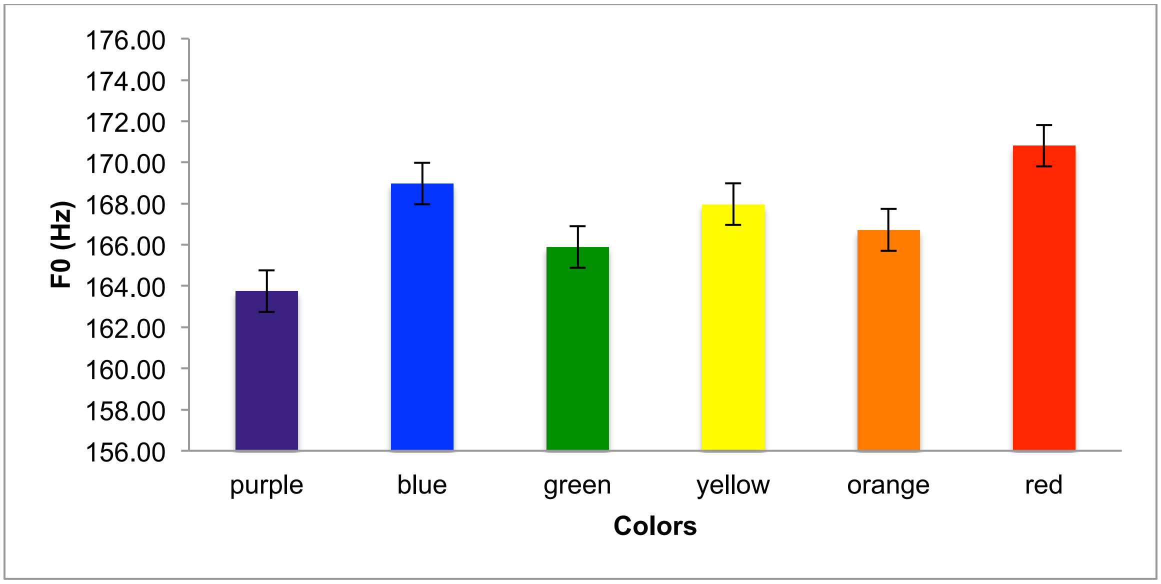


Figure 9. a) Average duration across brightness levels by blocks at word-level in Experiment 2. b) Average duration across brightness levels by blocks at sentence-level in Experiment 2. Error bars represent one standard error of the mean.

a)



b)

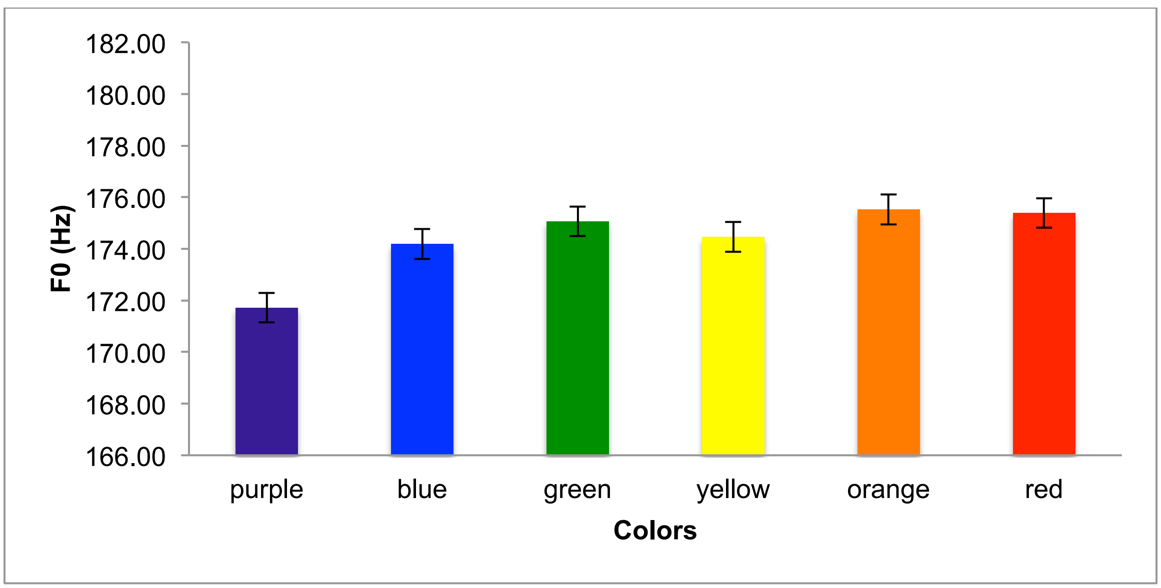
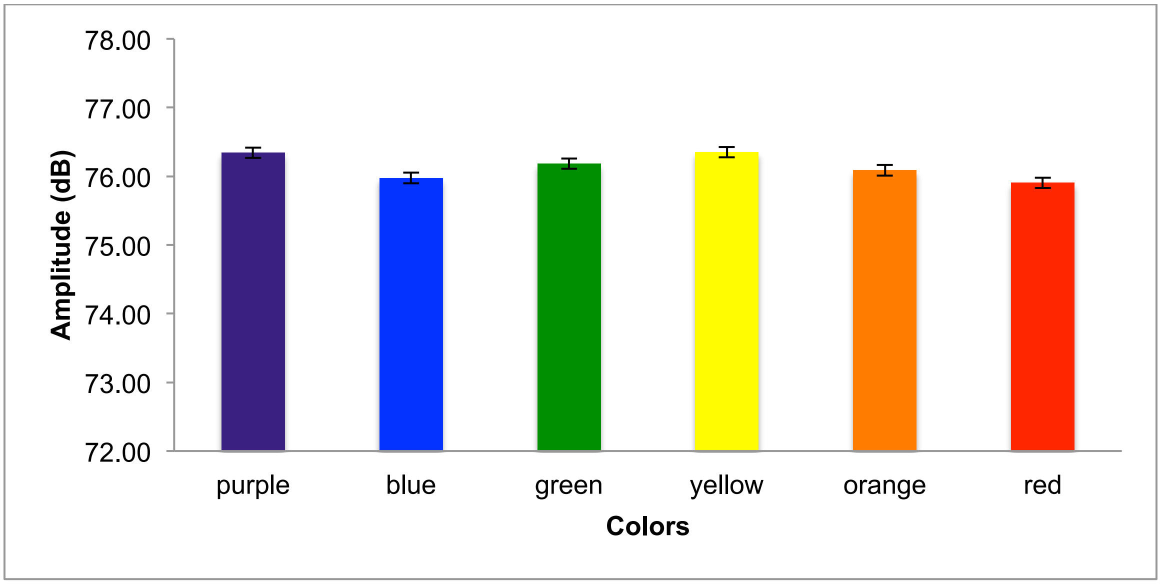


Figure 10. a) Mean F0 across colors at word-level in Experiment 2. b) Mean F0 across colors at sentence-level in Experiment 2. Error bars represent one standard error of the mean.

a)



b)

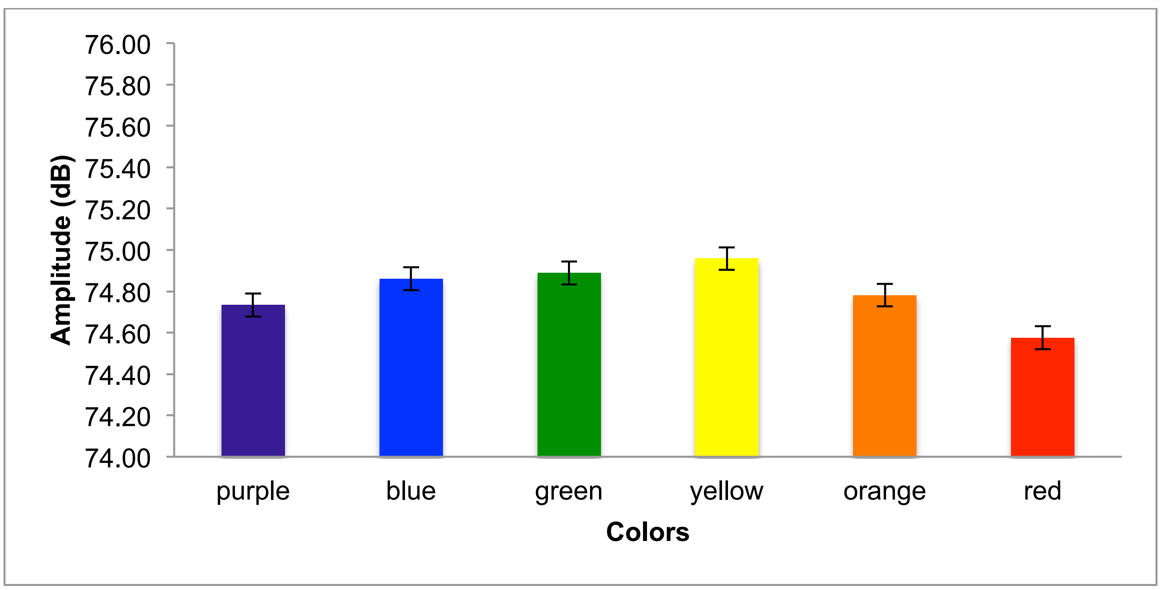
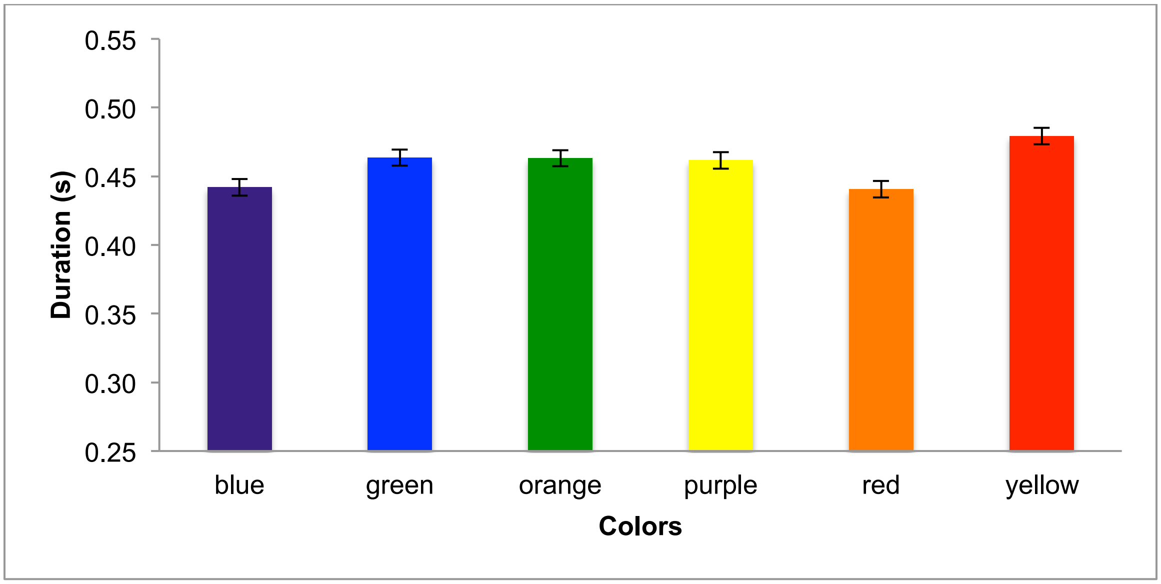


Figure 11. a) Mean amplitude across colors at word-level in Experiment 2. b) Mean amplitude across colors at sentence-level in Experiment 2. Error bars represent one standard error of the mean.

a)



b)

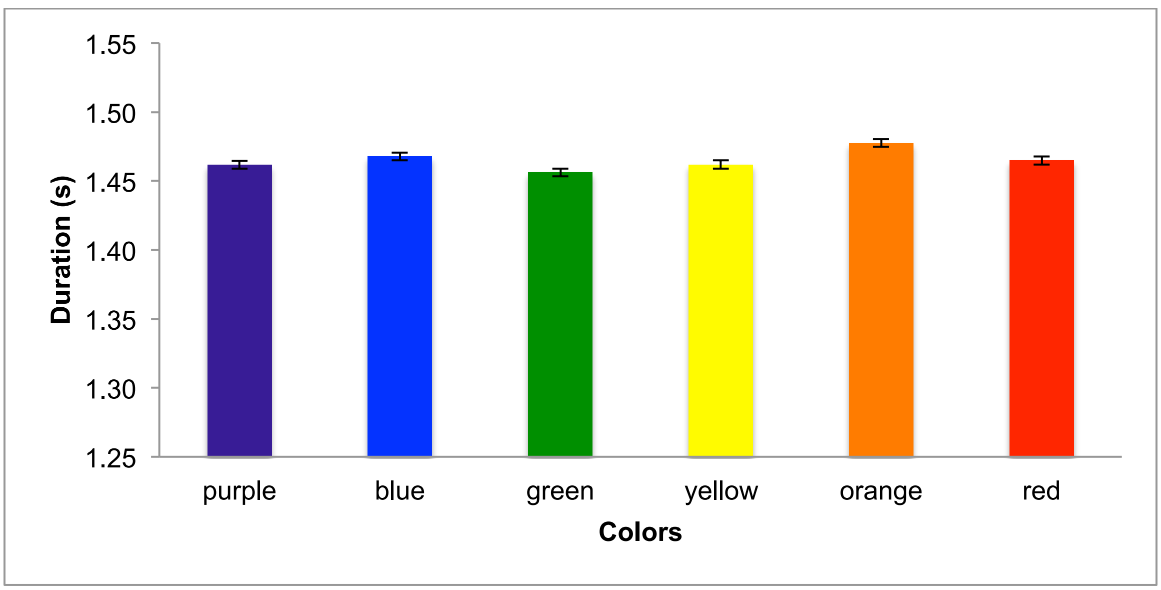


Figure 12. a) Average duration across colors at word-level in Experiment 2. b) Average duration across colors at sentence-level in Experiment 2. Error bars represent one standard error of the mean.

Appendix A: Color Spectrums

R: 0	R: 0	R: 0	R: 0	R: 0	R: 51	R: 102	R: 153	R: 204
G: 0	G: 0	G: 0	G: 0	G: 0	G: 51	G: 102	G: 153	G: 204
B: 51	B: 102	B: 153	B: 204	B: 255	B: 255	B: 255	B: 255	B: 255

R: 0	R: 0	R: 0	R: 0	R: 0	R: 51	R: 102	R: 153	R: 204
G: 51	G: 102	G: 153	G: 204	G: 255	G: 255	G: 255	G: 255	G: 255
B: 0	B: 0	B: 0	B: 0	B: 0	B: 51	B: 102	B: 153	B: 204

R: 51	R: 102	R: 153	R: 204	R: 255	R: 255	R: 255	R: 255	R: 255
G: 25	G: 51	G: 76	G: 102	G: 128	G: 153	G: 178	G: 204	G: 229
B: 0	B: 0	B: 0	B: 0	B: 0	B: 51	B: 102	B: 153	B: 204

R: 51	R: 102	R: 153	R: 204	R: 255	R: 255	R: 255	R: 255	R: 255
G: 51	G: 102	G: 153	G: 204	G: 255	G: 255	G: 255	G: 255	G: 255
B: 0	B: 0	B: 0	B: 0	B: 0	B: 80	B: 122	B: 153	B: 204

R: 51	R: 102	R: 153	R: 204	R: 255	R: 255	R: 255	R: 255	R: 255
G: 0	G: 0	G: 0	G: 0	G: 0	G: 51	G: 102	G: 153	G: 204
B: 0	B: 0	B: 0	B: 0	B: 0	B: 51	B: 102	B: 153	B: 204

R: 25	R: 51	R: 76	R: 102	R: 127	R: 153	R: 178	R: 204	R: 229
G: 0	G: 0	G: 0	G: 0	G: 0	G: 51	G: 102	G: 153	G: 204
B: 51	B: 102	B: 153	B: 204	B: 225	B: 255	B: 255	B: 255	B: 225

Note: During the experiments, participants did not see the RGB coordinates listed. The same set of spectrums was used in Experiment 2; however the labels were changed to novel words. Each color spectrums had three different versions (with the arrow pointing at the lightest shade, the darkest shade, and the intermediate shade). Only one version of each color is presented in the appendix.