2014

Visual Cognition While Listening to Stories

Nathalie Borges
Andrews University, borgesn@andrews.edu

J. N. Andrews Honors Program
Andrews University

HONS497
Honors Thesis

Visual Cognition While Listening to Stories

Nathalie Borges

April 8, 2014

Dr. Karl G. D. Bailey

Primary Advisor Signature:_____

Department:_____

Abstract

Different people have different perceptual patterns in response to language. Different language accents may influence visual cognition by increasing cognitive load when different objects are referenced in stories. The purpose of this study was to record differences in eye movement patterns while subjects listened to stories in two different accents, which served as the accent-induced stimulus load to increase cognitive processing. Using the eye tracker, eye movement patterns were recorded as subjects looked at an object array on a computer screen. My hypothesis that subjects would delay eye movement patterns to target objects while listening to stories in a foreign accent as opposed to an American accent because of increased cognitive load was supported with a main effect in accent type of stories.

Human Gaze Control in Response to Different Language Accents

While looking at a scene, people direct their eye movements toward targets that are relevant to their current task (Henderson, 2003). Top-down processes include prior knowledge, previous experiences with the scene, and goals and expectations. However, gaze control also involves bottom-up processes when eye movements are controlled by basic visual features such as color, size, and shape of an object. Both processes work together to direct eye movements throughout the scene.

The visual information in the scene thus plays an important role in determining where people will fixate and how long they will look at a particular location (Henderson, 2003). However, there are many other factors that contribute to scene perception besides visual stimuli including auditory and kinesthetic stimuli. For example, Huettig and Altmann (2005) observed eye movements in response to language involve an overlap between conceptual information conveyed by individual spoken words and conceptual knowledge associated with visual objects.

One of the earliest demonstrations of this phenomena involved short stories that subjects listened to while viewing a grid of nine simple line drawings (Cooper, 1974). While subjects listened to the story, their eyes fixated on certain objects that were related to certain key words in the story almost immediately after the word was uttered. For example, when subjects heard the word *lion*, they would immediately look at the picture of the lion. But is it always the case that listening to language will drive eye movements to related objects?

To answer this question, Cooper's (1974) study examined the extent to which the meaning of spoken language is able to continually control the locus of eye fixation. He observed that there were three types of fixation modes during scene perception in response to language: visual-aural interaction mode, free-scanning mode, and point-fixation mode. In the visual-aural

mode, subject's eyes fixated on targets related to words just after the words were uttered. However, in free-scanning mode, control of eye movements by the auditory stimulus was disengaged and subjects continually altered their gaze direction to words that were unrelated to the words being presented. Likewise, in point-fixation mode, subject's eyes fixated on one spot independent of the meaning of the utterances that they were hearing. Cooper (1974) observed that people frequently switched fixation modes while listening to spoken language. Cooper suggested that the different fixation modes reflected differences in the distribution of attention between auditory and visual stimuli, but did not propose a mechanism by which those differences might occur.

Nevertheless, these results are consistent with studies of gaze control that have been conducted since Cooper's pioneering work. Doherty-Sneddon, Bruce, Bonner, Longbotham, and Doyle (2002) observed that when people face high cognitive loads, they avert their gaze at critical points within a task or interaction to avoid processing of unnecessary, distracting, or arousing visual cues from their environment or to focus on task related comprehension. Similarly, Glenberg Schroeder, and Robertson (1998) suggested that averting eye gaze helps people to disengage from environmental stimulation, thereby enhancing the efficiency of internally directed cognitive processing. These results are consistent with the point-fixation mode of processing.

Ehrlichman and Micic (2012) explained that eye fixations wander off from targets when engaging in high cognitive load tasks. Tasks that involved minimal long-term memory search and required focus on material in working memory produced the lowest eye movement rates, whereas tasks with high requirements for long-term-memory search produced the highest eye movement rates. They observed that people make multiple eye movements unrelated to the

visual environment when they scan their surroundings for information in long-term memory and very few eye movements when they focus on information in the working memory buffer.

In addition to manipulation of auditory stimulus, Clarke and Garrett (2004) observed that native speakers tend to have difficulty processing language when listening to a foreign accent due to different pronunciation of words and syllable structure, but soon become accustomed to the foreign accent. They explained that accented speech can slow perceptual processing due to poor word identification, which relates to the purpose of foreign accents as the perceptual load in my research. When people process different accents, they might use the visual world to search for surrounding information in their environment to help them understand the accent, thus, delaying fixation onset to target objects.

Adank, Hagoort and Bekkerring (2010) have shown in their study that performance goes down when listening to foreign accents. They observed that people perform better in comprehension tasks when they imitate the designated foreign accent. Imitation may help subjects understand how non-native speakers pronounce words and syllables to better comprehend the foreign language to accomplish their designated tasks in the experiment. Likewise, if subjects imitate the accented speech in their minds, perhaps this will affect their eye movements.

This experiment examined whether cognitive load plays a role in the distribution of attention between auditory and visual stimuli by manipulating the load due to either auditory, visual, or extraneous tasks. This experiment attempted to replicate and extend the Cooper (1974) study by manipulating the auditory aspect of the experiment by instructing subjects to listen to stories in different accents of the English language. Specifically, I examined whether the time to fixate on language-related objects on a screen in response to target words in a story increases as

cognitive load is increased (due to shifts to point-fixation or free-fixation). I measured when subjects' eyes fixated on specific targets and where they looked when not focusing on the targets. When subjects engage in high cognitive load that includes multiple tasks, they may stop doing one task to better focus on another task, which may lead their eyes to look away from their immediate environment. Based on Cooper's (1974) study, working memory load would be expected to influence point fixations, but would perceptual load such as foreign accents have the same effect? Do different types of cognitive loads lead to different off-task behaviors? The purpose of this study was to record differences in eye movement patterns while subjects listened to stories in two different accents, which served as the accent-induced stimulus load to increase cognitive processing. In my study, I hoped to find whether different loads lead to more or less of task behavior and try to identify what kind of off-task behavior is occurring.

Methods

Subjects: This study consisted of a convenience sample of 26 self-selected students from the Behavioral Sciences Research Participant Pool. Four of the subjects had to be removed from the sample because of invalid data from eye tracker malfunction or too much blinking, which makes recording eye movement patterns more difficult. The participants spoke American English as their first language. There was no preference of ethnicities among subjects. Subjects received research credit for participation.

Materials: Visual stimulus materials consist of 12 slides on a computer screen of a 3 X 3 matrix of nine distinct objects each that are related to key words uttered in the stories. Visual images were all in color. The auditory stimulus consisted of 12 short stories, six for each accent, uttered in two different accents: Standard American and Caribbean (Trinidad and Tobago). The Standard American accent served as the control accent. Each story lasted approximately two to

three minutes. Eye movements were recorded using a 60Hz Arrington Viewpoint eye tracker, which records length of fixation, frequency of fixation, and the time it takes for the eyes to fixate on targets. Each subject took approximately 30 minutes to complete the task, including eye tracking setup.

Procedure: Subjects were instructed to listen to 12 stories while leisurely looking at images on a computer screen. The stories were in two different accents: Standard American and Caribbean. These stories were counter-balanced to eliminate stimulus effects. Cognitive load was manipulated by changing the auditory stimulus to a Caribbean accent, which served as the unfamiliar accent to increase cognitive load. The unfamiliar accent condition provided the cognitive load condition. Subjects were also instructed to listen carefully to each story because they would be asked a question regarding the content of each story after each story was heard. This short quiz after each story served as the working memory load task. The comparison of eye movement patterns during accented and non-accented stories while paying attention to the stories allow for comparison between the effects of perceptual auditory load on eye movements and the effects of working memory load.

Subjects were hooked up to the eye tracking device throughout the experiment. Subjects looked at different objects on a computer screen while listening to a story in a different accent. Subjects were only instructed to look freely at the screen during the experiment because further instructions may have inflicted experimenter bias and subject bias. My independent variables was the auditory stimulus of language accents and the type of reference to target objects on the screen (direct or indirect). My dependent variable was the length of time to first fixation and the frequency of fixations on a specific target on the screen.

Analysis: The data was analyzed into two steps. The first step of data analysis was using two within-subjects ANOVAs to examine trial-level fixation patterns. One test analyzed the number of fixations by accent, while the other test analyzed the length of fixations by accent. The second part of data analysis was also using a 2-way within-subjects ANOVA to examine word level effects with four conditions: accent stories with direct words, accented stories with indirect words, non-accented stories with direct words, and non-accented stories with indirect words.

Results

After testing for average length of fixation and average frequency of fixation during the experiment, there was no significant difference of eye movement patterns between native and non-native accents. For average length of fixations: $F(1,21) = 0.459$, $p = .51$; for average frequency of fixations: $F(1,21) = 0.596$, $p = .60$. Being that my p-values for both average length of fixation and average frequency of fixation were larger than 0.05, these results were not significant and failed to reject my hypothesis that there would be a difference in eye movement patterns between accented and non-accented stories.

In terms of eye movement patterns on the computer screen, I noticed that subjects, on average, had the tendency to stare at the center of the screen, similar to Cooper's (1974) findings. According to Figure 2A. And Figure 2B., subjects looked at the center of the screen for a majority of the experiment both for the Standard American accented stories and the Caribbean accented stories. For the purpose of this study, I only analyzed the first minute on average of each story for all the subjects to formulate the graphs (Figure 1A. and Figure 1B.). Although there was a slight change in eye movement frequency for both accented stories, there was not enough significant change to reject my hypothesis. According to Figure 2A. and Figure 2B.,

accented stories caused a slight difference in the frequency of eye movement fixations to certain objects on the screen because subjects tended to look at the center of the screen for both accents, but these results were later proved significant on a word-level analysis.

After further analysis of target objects referenced both directly and indirectly in both native and non-native accented stories, Figure 3A. and Figure 3B. showed the fixation patterns to target objects and non-target objects on the computer screen. In Figure 3A., the graph shows the fixation to target objects in time frames at 60 frames per second. Subjects had a higher proportion of fixation on target objects during accented stories for both direct and indirect object references. In Figure 3B., the graph shows fixation to other objects in the display in time frames at 60 frames per second. Subjects tend to have a higher proportion of fixation to other objects on the screen when target objects are referenced indirectly in both accented and non-accented stories.

To test for significant change in word-for-word analysis of fixation patterns to target objects and non-target objects, I used two 2 x 2 within-subjects ANOVA. One test analyzed fixations to target objects, the other test analyzed fixations to non-target objects. My two conditions were accent versus non-accent stories and indirect versus direct object reference. There was no interaction between accent type and object reference type ($F(1,21) = 0.141$, p-value $= 0.711$). There was a marginal difference in indirect versus direct object reference ($F(1,21) = 3.149$, $p = .091$). There was a significant main effect in comparing accent versus non-accent stories ($F(1,21) = 10.13$, $p = 0.0045$).

Discussion

Although I did not find any significant difference of average frequency of eye fixations and average length of fixations between accented and non-accented stories for the first

60 seconds of each trial, there was a main effect in accent type in word-level analysis, which indicates that there was a difference in fixation patterns between native accented and non-native accented stories. There was no difference in accent type in trial-level analysis, but there was a difference in word-level analysis. This implies that non-native accents made a difference in fixation patterns on a word-for-word level when subjects listened to the stories and observed target and non-target objects on the screen. Although there was no interaction between accent type and reference type, I still found a main effect that explained some fixation differences on Figure 3A. and Figure 3B.

I did find that subjects had the same tendency to stare at the center of the screen similar to Cooper's (1974) study, regardless of auditory stimulus manipulation (Figure 2A. and Figure 2B). However, I only analyzed the average number of fixations for all subjects within the first minute of each story for the course of the experiment. Each story had specific target words that were either referenced directly or indirectly. This served the purpose to track eye movement patterns to specific objects on the computer screen when they were referenced in the stories. The tendency for subjects to stare at the center of the screen reflects Cooper's (1974) explanation of the point-fixation mode, where subjects focused on one spot of the screen, independent of what they were hearing.

I noticed during the course of the experiment, many subjects had the tendency to close their eyes and slow their body movements. This may be due to eye muscle fatigue from staring at the computer screen for a long period of time or the students may have already been tired prior to the experiment. Subjects had a hard time keeping their eyes open around the eighth or ninth story. This may have affected my results slightly when comparing average frequency of fixations and average length of fixations.

According to Figure 1A. and Figure 1B., there was no significant difference in trial-level fixation patterns between native and non-native accented stories. However, there were effects in word-level fixation patterns. This indicates that perhaps the auditory stimulus was not as effective, for a whole trial instead of word for word analysis, as in other studies such as Clarke and Garrett's study in 2004. My choice of using a Caribbean accent as my experimental variable may not have been as effective as using perhaps a French or Spanish accent. Different foreign accents might have different effects on working memory load tasks for different people. However, the subjects in my study attend a richly diverse campus, with students from all over the world who have many different accents. Students come in contact with each other every day and they soon become accustomed to the diversity of foreign accents. This in turn may have had an effect on eye movement patterns throughout both accented and non-accented stories. Perhaps if this study had been conducted on another campus or to a specific ethnic group, my results could have been different.

After further analysis of word for word onset indicating target objects on the screen, Figure 3A. and Figure 3B. showed the fixation patterns to target objects and non-target objects, respectively. In Figure 3A., subjects had a higher proportion of fixation to target objects on specific regions of interest on the computer screen when the target objects were referenced both directly and indirectly in accented stories. This reflects Cooper's (1974) explanation of the visual-aural fixation mode, where subjects directed their gaze to target objects right after they were referenced in the stories. In this case, subjects had a higher probability of looking at target objects when referenced in the foreign accent. This indicates that subjects focus more on target objects when listening to a foreign accent, which supports Huettig and Altmann's (2005) observations of fixation patterns in response to language involve an overlap between conceptual

information conveyed by individual spoken words and conceptual knowledge associated with visual objects. These results could be because subjects are forced to pay more attention when they hear words in a foreign accent, directing their gaze to target objects referenced in the stories. Subjects focused more on target words when hearing the accented stories whether the target objects were referenced directly or indirectly.

In Figure 3B., subjects had a higher proportion of fixations to other objects when references were made indirectly for both accented and non-accented stories. This reflects Cooper's (1974) explanation of the free-scanning mode, where subjects look at objects that are unrelated to target objects uttered in the stories. This can indicate that subjects have a tendency to search for other objects that give any information related to words that are referenced indirectly. With indirect object references, subjects look around at other objects to make sense of their scene perception that can help direct their gaze to the target object referenced in the story. Figure 3B. showed that subjects tend to look more at other objects when target objects were referenced indirectly throughout the course of the task.

Unfortunately, there exists no perfect research study. Each subject took approximately 30 minutes to complete the task, which fatigues subjects' eyes and leads them to fixate on objects less than expected. Fixation patterns could have been skewed by subjects failing to keep their eyes open during the task. Another factor that could have affected the results was the variety of ethnicities of the subjects. Different ethnicities have different accents, making it difficult to single out the changes in fixations from the Caribbean accented stories. Andrews University is a universally diverse school that consists of students from all around the world. On a daily basis, students interact with many people from different countries with different accents, which

eventually tune out the difficulties in language comprehension, as noted in Clark and Garrett's (2004) study of rapid adaptation to foreign-accented English.

This research study indicates that foreign accents may have an influence on how people perceive things in scene perception. As people hear references and such in a foreign accent, they tend to look to surrounding scenes to gather information that may help them better understand what is being presented to them. Language plays an influential role during scene perception because language helps people direct their gaze to things that are relevant to a specific task. Language gives meaning to a scene or specific object, which directs people's attention to what needs to be seen and recognized to gain a better understanding of one's surroundings. Along with language, people use sensations and past experiences to put pieces together to form a bigger picture of what one perceives. This study shows that language does have an influence on how people perceive things, which can help people communicate with each other more efficiently. Also, my findings confirm that subjects use the visual world to help interpret language when listening to a different accent, as shown in Figure 3B.

References

Adank, P., Hagoort, P., & Bekkerring, H. (2010). Imitation Improves Language Comprehension. *Psychological Science*, *21*(12), 1903-1909.

Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *Acoustical Society of America*, *116*(6), 3647-3658.

Cooper, R. (1974). The Control of Eye Fixation by the Meaning of Spoken Language. *Cognitive Psychology*, *6*, 84-107.

Doherty-Sneddon, G., Bruce, V., Bonner, L., Longbotham, S., & Doyle, C. (2002). Development of Gaze Aversion as Disengagement. *Developmental Psychology*, *38*(3), 438-445.

Ehrlichman, H. & Micic, D. (2012). Why do people move their eyes when they think? Current Directions in *Psychological Science, 21,* 96-100.

Glenberg, A. M., Schroeder, J. L, & Robertson, D. A. (1998). Averting the gaze disengages the environment and facilitates remembering. *Memory and Cognition, 26,*(4), 651–658.

Henderson, J. M. (2003). Human Gaze Control During Real-World Scene Perception. *Trends in Cognitive Sciences*, *7*(11), 498-504.

Huettig, F., & Altmann, G. (2005). Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. *Cognition*, *96*, B23-B32.

Figures 1A. and 1B.

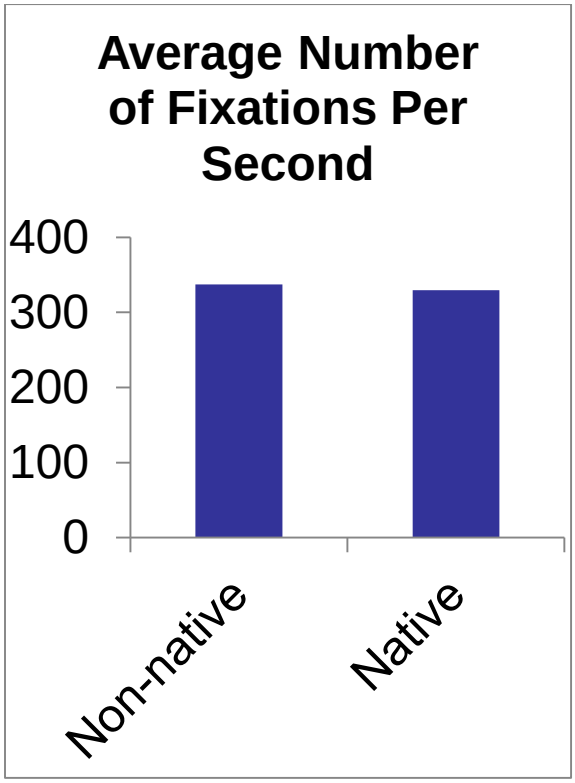

Figure 1A. shows average number of fixation per second (for the first 60 seconds of each story) between native and non-native accented stories.

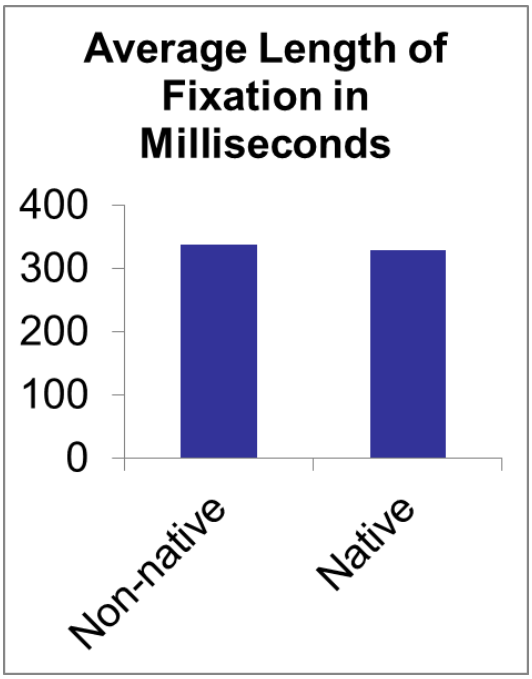

Figure 1B. shows average length of fixation in milliseconds (for the first 60 seconds of each story) between native and non-native accented stories.

Figures 2A. and 2B.

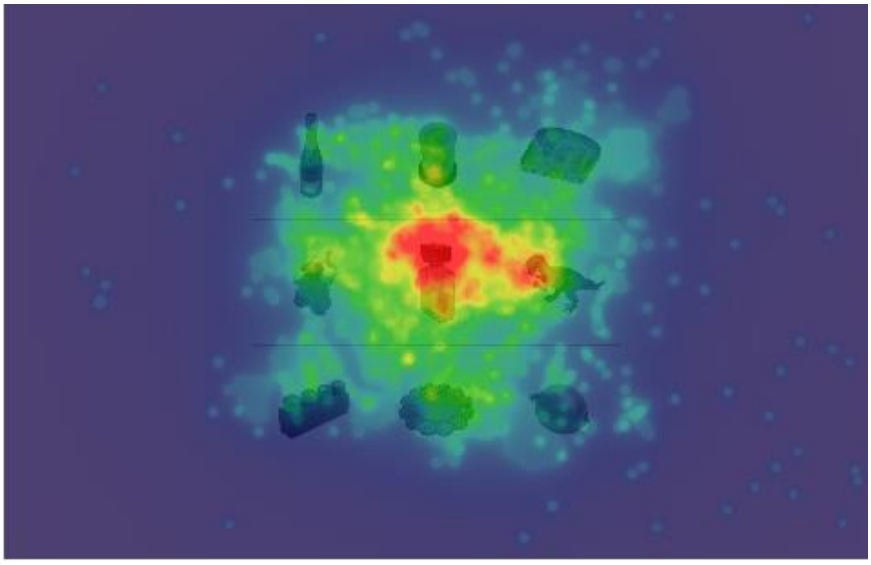

Figure 2A. Probability of fixation on different parts of the display for native accented stories.

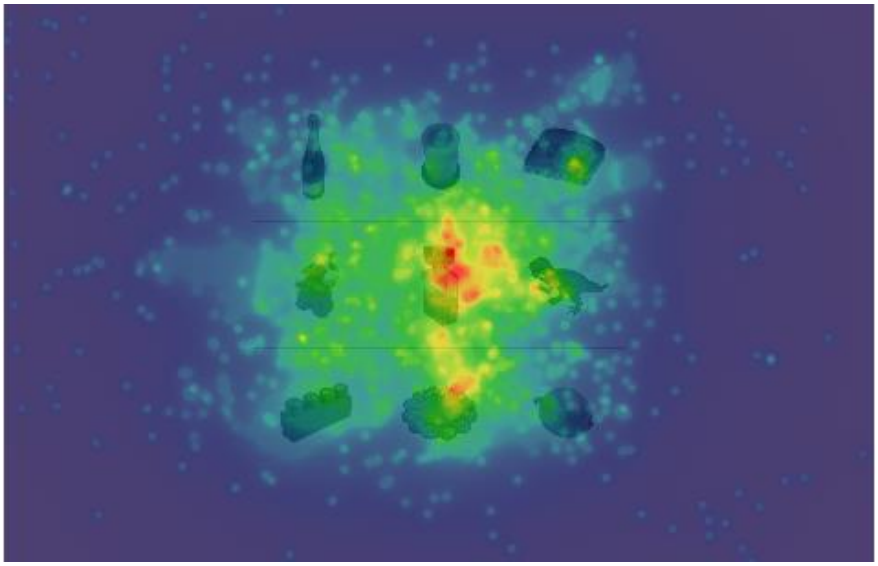

Figure 2B. probability of fixation on different parts of the display for non-native accented stories.

Figure 3A. and Figure 3B.

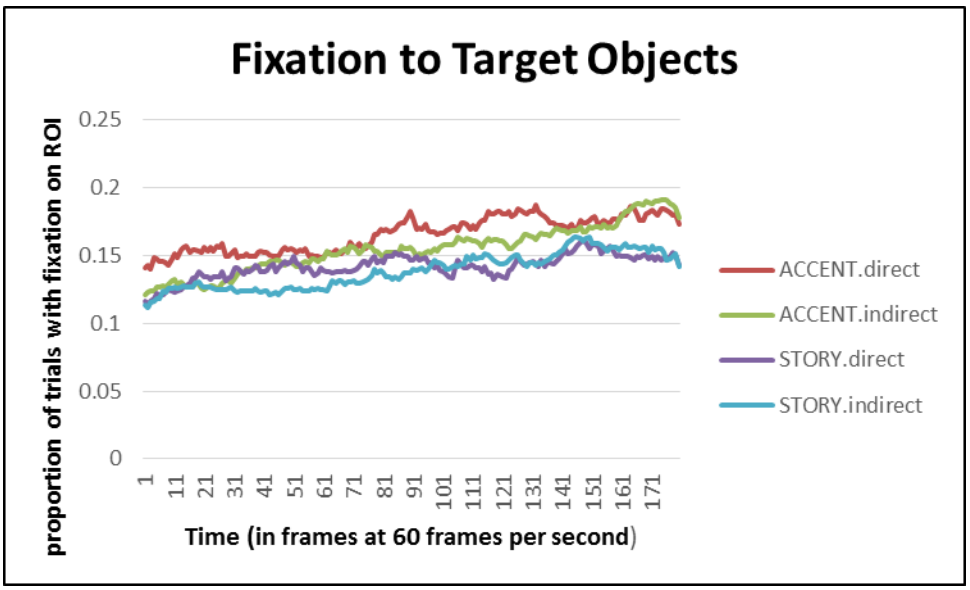

Figure 3A. Fixation to target objects in time frames at 60 frames per second. Subjects had a higher proportion of fixation on target objects during accented stories for both direct and indirect object references.

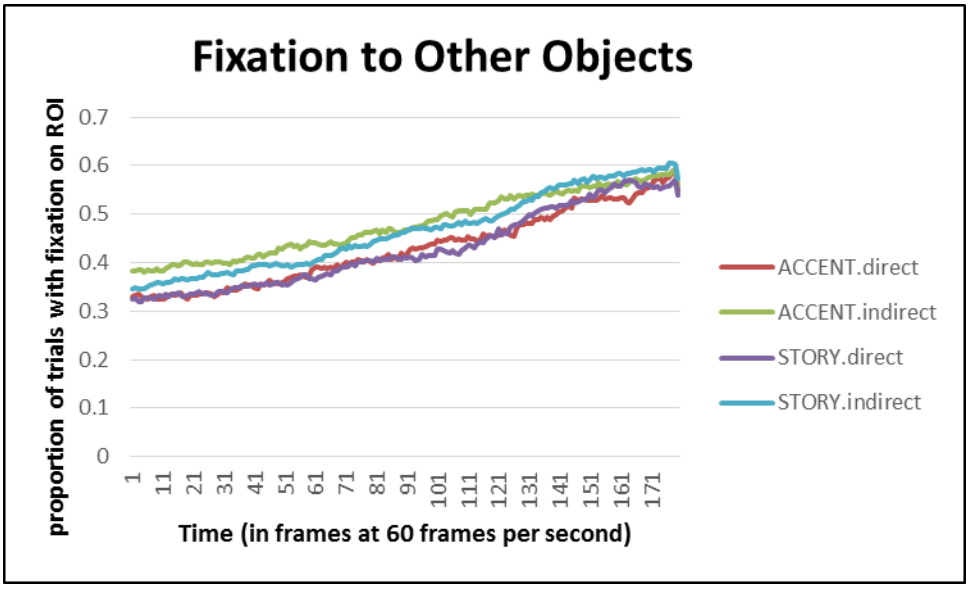

Figure 3B. Fixation to other objects in the display in time frames at 60 frames per second. Subjects tend to have a higher proportion of fixation to other objects on the screen when target objects are referenced indirectly in both accented and non-accented stories.