How Our Thoughts About Language Influence What We Hear

Antoinette Watson[1,2], Ebony Pearson[1], and Duane G. Watson[1]

[1]Department of Psychology & Human Development, Vanderbilt University, Nashville, TN, 37203
[2]The School for Science and Math at Vanderbilt, Nashville, TN 37203

BRIEF. In this study, we wanted to conduct a statistical analysis on a set of questions concerning a person's accent.

ABSTRACT. Social environment and geographical location often influence a person's dialect. Pronunciation-specific vowels can be an indicator of their regional background and upbringing. The pin-pen merger is where the vowel sounds /i/ and /e/ merge before nasal consonants like /n/ and /m/. People with this merger would pronounce words such as "gym" and "gem" similarly, while those without the merger can differentiate. This study used a visual world eye-tracking paradigm with 71 participants. Twenty-eight were merged speakers, and 46 were non-merged speakers. Participants selected the image for each word they heard. The task involved choosing a target image (pin words) and not the distractor image (pen words). When participants heard the word "pin," they saw four images, a sewing pin, a writing pen, and two distractor images. In a survey, 28 respondents believed they had an accent, while 46 did not. Accented participants were expected to be merged, while those nonaccented would not. Individuals with accents are perceived as merged speakers, whereas individuals without accents are the opposite. This research helps researchers understand these influences and helps us better understand how language works in our minds and can help us understand how different dialects and accents affect communication.

INTRODUCTION.

In linguistics, language variation refers to how individuals or groups of speakers produce spoken or written language differently based on various factors. This variation can manifest in several ways, including pronunciation, word choice, grammatical structures, and speech rate. Language perception refers to listening, while language production refers to speaking.

The pin-pen merger is a prominent feature in the southern region of the United States. The pin-pen merger is a conditional merger where the vowel sounds /i/ and /e/ merge before nasal consonants (i.e., the sounds /n/ and /m/.) [1] For example, Pin and Pin are pronounced similarly. The idea of merged speakers are speakers that pronounce two similar words, with two different spellings, similarly. In contrast, non-merged speakers pronounce two similar words, with two different spellings, differently.

Dialect is the influence of a person's region or social group on language. For example, African American English (AAE) is a type of English that includes grammatical features, syntax, and word usage. At times, African Americans can speak AAE in formal speech, informal speech, and culturally influenced speech. [2] Another example is Mainstream American English. A person's dialect is the reason their language background and experience are the influences on why and how they pronounce words.

Speech production and perception, the pin-pen merger, and dialect all have a connection to the way language is processed in someone's mind. These specific topics have a direct correlation to each other because they all impact how we hear and interpret specific words within the pin-pen merger and how we understand the words being pronounced. Since our dialects are influenced by our social groups and where we live, it can impact the way we pronounce and communicate specific words and phrases. [1]

In this study, we wanted to determine whether speakers with accents were more likely to be merged speakers and how being a merged speaker might impact visual language processing. We hypothesized that speakers with accents are more likely to be merged speakers. Additionally, we assumed that speakers with an accent would spend more time looking between the two words "Pin" and "Pen", which would indicate that they are merged speakers. This was done by doing an eye-tracking experiment. Overall, we expected that this research may shed light on how communication impacts day-to-day conversations with people and how pronunciation impacts how we hear, interpret, and understand language.

METHODS.

Methods.

In this experiment, we wanted to conduct a statistical analysis on a set of questions that are asked post-experiment. We surveyed a subset of questions to determine if participants spoke with or without accents, which allowed us to determine whether a speaker was considered merged or non-merged.

Demographics Survey.

To complete the questionnaire, we recruited seventy-one participants from a research university, aged eighteen and older; in which all participants were required to complete informed consent, allowing us to review all data provided. Institutional Review Board approval was also obtained to ensure all proper ethics, rules, and guidelines were followed. Participants identified as White (59.6%), Asian (22.5%), African American (12.4%), Native Hawaiian (1.1%), American Indian/Alaska Native (1.1%), or chose not to specify (3.4%). Participants resided in various states. During this experiment, all participants (N=71), were classified as either merged or non-merged speakers based on the vowel production of words in the pin-pen merger. Before starting the experiment, participants were asked to complete a demographic survey called the Language and Social Background Questionnaire. The questionnaire reports information on the participant's demographic information, education level, and native language. [3]

Stimuli Training.

Before completing the eye-tracking task, participants were required to complete a stimulus training session through a picture naming task implemented on Gorilla Experiment Builder. In this session, participants were seated at a computer for the learning task and presented images one by one with their corresponding names. After viewing each image, participants were presented with each image again and prompted to type the word that corresponded to each image from the previous learning task. If a participant entered an incorrect response, the image was returned to the group of remaining images to be shown again until the participant achieved a correct response and all items had been answered accurately. [1]

Eye-Tracking Experiment.

In this experiment, we utilized an EyeLink 1000 Plus eye-tracker. [4]Once participants completed the stimuli training, they progressed to the visual world paradigm eye-tracking task. The paradigm of the experiment is a research method used in psycholinguistics to study how individuals process language in real-time during activities like listening

Table 1. This is the wordlist from which participants are asked to read. These words are put into a random order and participants are asked to read down the list of words, as shown above.

PEN	PIN	BET	BIT	FILLERS		
pen	meant	ben	ten	cents	Jen	Ken
pin	mint	bin	did	since	gin	kin
pet	dead	bet	head	set	bed	keds
pit	did	bit	hid	sit	bid	kit
cat	bat	dog	talking	task	water	monster
ham	pal	robin	cow	front	frog	had/ carrot

and reading. The researchers observe participants' eye movements while listening to words and looking at pictures. Researchers review and observe eye movement patterns and make predictions about the cognitive processing occurring while the participants hear words and see various images. Understanding this helps researchers determine how we process language and information from the world around us. Each set of images includes a target image, a competitor image, and two phonetically unrelated distractor images. This design allowed us to examine how the participants reacted when they heard a specific word and saw related images in the context of the pin-pen merger. The participants were presented with the four images followed by hearing a target word, which always consisted of the "pin" production of the vowel. The participant was instructed to click on the image that best matched the word that they heard. After the participant selected the target word, the screen would generate the next set of images. The eye tracker gathered participant fixation points and eye movements during the experiment. We conducted this experiment to observe the difference, if any, in language processing of merged and non-merged speakers. For the merged participants, the pronunciation of "pin" is expected to show more fluctuation in their eye movements when presented with the target competitor (pin) and phonological competitor (pen). Participants might spend more time processing both words before making a choice. For non-merged participants, the pronunciation of "pin" is predicted to exhibit more distinct eye movements to the target item. Non-merged participants are expected to take less time fixating on the target because they maintain the phonemic distinction between "pin" and "pen." Their eye movements are likely to indicate a more efficient and accurate language input processing.

Production Task.

We prompted participants to read a wordlist and a brief paragraph to determine their mergedness. The wordlist was a standardized practice that the researchers utilized in past experiments. [3] The wordlist included pin-pen words with /i/ and /e/ sounds followed by nasal (e.g., pen/pin) as well as distractor words with //i/ and /e/ sounds followed by non-nasals (e.g., pet/pit) to gather a baseline of their production of the sounds. The paragraph also included embedded pin-pen words to gauge production in continuous speech. Participants were seated in a quiet room with a computer and a microphone. They were presented with the wordlists and paragraphs on the computer screen and instructed to read them aloud clearly and naturally (Table 1)

Acoustic Analysis.

Following the recording of participants' readings, an acoustic analysis was conducted using Praat software. This analysis focused on extracting acoustic features, including vowel formant frequencies, to assess the extent of the pin-pen merger in participants' speech. The acoustic analysis involved measurements of the first and second formants (F1 and F2) for each vowel sound in the recorded speech samples. Formant values were used to quantify the spectral characteristics of the vowel

sounds, providing a quantitative measure of the pin-pen merger. The wordlist included pin-pen words (target competitor words and phonological competitor words) and distractors words.

Minimal Pair Task.

In this section of the experiment, participants read pairs of words and assessed whether they rhymed or not. We also used filler words in this section to increase the difficulty of the questionnaire. Participants observed two words simultaneously (i.e., "gem" and "gym" or "meant" and "mint") and were given the direction to identify which words rhymed when spoken aloud. To increase the difficulty, participants were asked if words like "monster" and "water" rhymed. This part of the experiment was conducted to determine if speakers with the merger are expected to believe that words like "gem" and "gym" rhyme, whereas speakers without the merger are expected to believe that the two words do not rhyme. [1]

Language Attitudes & Beliefs Survey.

A post-experimental questionnaire related to speech pronunciation and accent was collected at the end of the experimental procedure to determine if a participant was a merged or non-merged speaker.

RESULTS.

Eye-Tracking Experiment.

The participant's eye movements were measured during the eye-tracking experiment while they performed the visual world paradigm task. When collapsing the data between participants who self-identified as speaking with an accent versus not speaking with an accent, we see that accented participants experience an increased proportion of looks to the phonological competitor. This result indicates that accented speakers experienced increased difficulty after hearing the target word and attempting to select its matching lexical item. Additionally, the unaccented participants experience an increased amount of looks at the target item and competitor compared to the unrelated distractor items (Figure 1).

When observing the fixation proportions among all participants within the context of the pin-pen merger, a distinct pattern emerges. Participants exhibit a heightened proportion of fixations directed towards both the target word (e.g., "pin") and its phonological competitor (e.g., "pen") compared to the unrelated distractors. Notably, there are significant differences in fixation distribution within the initial 200-400 milliseconds duration, which indicates a shared allocation of attention between the target and its phonological competitor. Beyond the 500-millisecond mark, a noticeable shift is observed, with participants predominantly fixating on and selecting the intended target item.
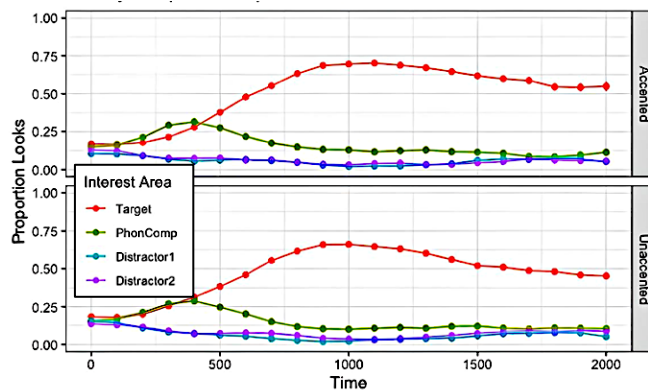

Figure 1. Survey Response: I speak with an accent. The graph shows that there is difficulty in choosing the target word when there is a phonological competitor present. There are fixation points before the target item is chosen.
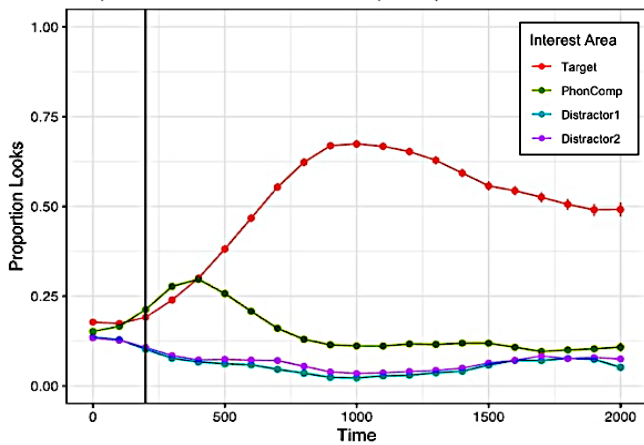
Figure 2. Proportion of fixations across all participants. The graphs show that there is difficulty in choosing the target word when there is a phonological competitor present. We see a fixation on the competitor before the participants choose the target item.

The early-stage fixation differences suggest a rapid assessment of phonological features, perhaps reflecting an initial ambiguity resolution process. Subsequently, the shift towards prolonged fixations on the target item implies a resolution in favor of the intended word. This dynamic interplay in fixation distribution sheds light on the cognitive mechanisms underlying the processing of phonological competitors, providing valuable insights into the intricate processes involved in resolving lexical ambiguity. [Figure 2.]

Production Task.

During the production task, participants were requested to read a short story and a wordlist this task helped determine the mergedness of the participant. For the merged participants, there was a rise in fixation points and reaction time. Before 500 milliseconds, merged participants pronounced the phonological competitor at first, and after 1000 milliseconds merged participants finalized the word that they said before moving on to saying the next word. Non-merged participants were able to pronounce the target word clearly and more quickly. After 500 milliseconds, non-merged participants realized both words and after 1000 milliseconds pronounced their target word.

Minimal Pair Task.

In the Minimal Pair Task, participants who selected "Yes" for questions that stated "gem and gym" or "mint and meant" rhymed were perceived as merged speakers and speakers who spoke with a Southern or Northern accent. Whereas participants who selected "No" for the questions as stated before, were perceived as non-merged speakers, and speakers who spoke without an accent.

Language Attitudes & Beliefs Survey.

Post-experiments, we found that participants who spoke with the merger had difficulty comprehending the pin-pen words. We also found that participants who did not speak with Merger needed help understanding the pin-pen words. Individuals who speak with an accent, such as a Southern accent or a strong accent are perceived to speak with the merger. Whereas, individuals who do not speak with an accent, such as mainstream English speakers, are not perceived to speak with the merger. 28 participants believed they spoke with an accent, while 46 participants believed that they spoke without an accent ($p=0.04$), which means there was a significant difference between speakers with an accent and without one.
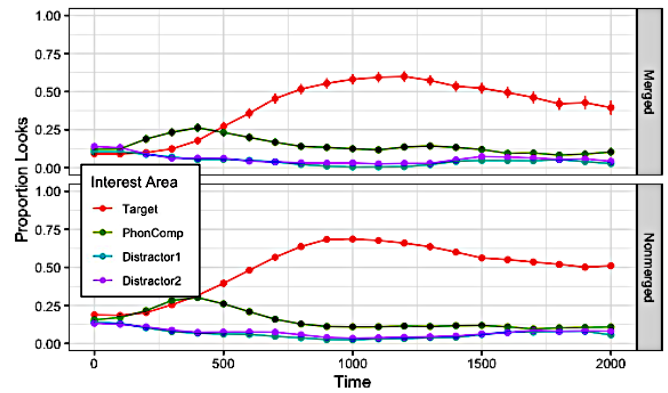

Figure 3. Mergedness: Production. Target is the /i/ words and distractors are the /e/ words. Participants hear the word "Pin" and then they see the four images, pin, pen, and two distractors. People with an accent tend to have more difficulty selecting the target words.

DISCUSSION.

In conclusion, speakers that were considered merged speakers had slower reaction times during the eye-tracking experiment. This is because the participants had a harder time differentiating between the words, which caused participants' brains to take a longer time to process the words heard and took longer to look between the pin-pen words. Speakers considered non-merged speakers had a faster reaction time during the eye-tracking experiment. This is because the participants had an easier time differentiating between the words, which caused participants' brains to process the words heard faster, and the participants looked at the answer.

Based on our sample, speakers who live in the south were considered merged speakers, and those speakers had a southern accent or spoke AAE. [2] This meant that those speakers had a harder time differentiating between the pin-pen words and had a slower reaction time. Speakers that were considered non-merged speakers either did not have an accent or spoke Mainstream English and could differentiate between pin-pen words and had a faster reaction time. Researchers typically observe speakers who live in the South as speakers to have the merger and wanted to investigate how a speaker's brain processes words.

If there had been more time, we would have been able to dig deeper into the results and how the results impact the speakers with accents. Also, we would have recruited more participants from the Midwest.

Most of the participants came from the Northern or Southern parts of the United States. Having some participants from the Northwestern region would be interesting data to investigate and determine a participant's accent and if they are a merged or non-merged speaker.

For further research, we want to conduct a separate research experiment that investigates a listener's expectations about the mergedness of a speaker given their dialect. The research question that we want to conduct is "Do listeners make assumptions about a speaker's mergedness given their dialect?" In the Production task, pre-training will be conducted to gather baseline and mergedness. In the Perception task, pre-training will be conducted to gather baseline including the /ɪ/ sound and the /ɛ/ sound to determine if they can detect a distinction between the phonemes. Participants will type the word that they hear, and this will determine if the person is a merged or non-merged speaker in perception.

While conducting the Language Attitudes and Beliefs survey, we concluded that people's beliefs and attitudes about language can impact

how a person understands what they hear. Understanding these influences can help researchers better understand how language works in our minds. Finally, research can help us better understand how different dialects and accents affect communication.

REFERENCES.

[1] Austen, M. Production and perception of the Pin-Pen merger. *Journal of Linguistic Geography*, **8**(2), 115–126 (2020).

[2] King, S. From African American Vernacular English to African American Language: Rethinking the Study of Race and Language in African Americans' Speech. *Annual Review of Linguistics,* **6**, 285-300 (2020)

[3] Anderson, J., Mak, L., Chahi, A. K., & Bialystok, E. The language and social background questionnaire: Assessing degree of bilingualism in a diverse population. *Behavioral Research*, **50**, 250-263 (2018).

[4] Kiwako, & Campbell-Kibler. SPEAKER-ADAPTATION TO /ɪ/ - /ɛ/ MERGER: AN EYE-TRACKING STUDY. *17ᵗʰ International Congress of Phonetic Sciences,* 954-957 (2011).

Antoinette Watson is a student at Martin Luther King Jr. Academic Magnet High School in Nashville, TN; she participated in a research internship through the School for Science and Math at Vanderbilt.