University of Utah
UNDERGRADUATE RESEARCH JOURNAL

INVESTIGATING THE DIFFERENCE IN LISTENING EFFORT BETWEEN AMERICAN AND KOREAN LISTENERS FOR AMERICAN ACCENTED ENGLISH

Kalista Vordos (Carina Pals, Dahyun Kim, Adeline Han)
Department of Psychology

Abstract

This study aimed to explore the effects of English language processing on the cognitive effort required to comprehend spoken language, referred to as 'listening effort,' in native versus non-native speaker populations. Our experiment explored whether exposure to America-accented English affects the ability to benefit from constrained sentences versus unconstrained sentences. A convenience sample of 60 students across three global campuses in two countries (United States and South Korea) was recruited to participate. A fixed linear regression model showed a significant difference between the United States and the South Korean population regarding the effect of context on processing speed. These results indicate that there could be a significant difference in language processing between native and non-native speakers.

Introduction

When we listen, our comprehension depends on cognitive resources such as attention, processing speed, working memory, long-term memory, linguistic ability, and the difficulty of the speech signal. This, in turn, affects how much of those resources are needed for understanding the sound. In other words, the more demand for those cognitive resources a listener has, the more listening effort is required to comprehend auditory language.

Since every listener has different demands of cognitive resources, the listening effort dramatically varies between individuals. A minimal effort for speech understanding is needed if there is an easy match between language input and long-term memory (Rönnberg, Stenfelt, & Lyxell, 2013). However, explicit processing and further cognitive resources are required for speech understanding if the input does not easily match memory. This scenario is commonly seen with listeners of a non-native language or accent (Oosthuizen, Picou, Pottas, Myburgh, & Swanepoel, 2019). One frequent challenge that shows input-memory mismatch is the comprehension of a non-native language because the listener's mental representations have not yet been fully developed. People who are continually exposed to an accented language, whether or not they are native speakers of that language, should be able to decipher a particular accent faster than those exposed to an unfamiliar accent.

Listening effort can be measured in various ways, including pupillometry, dual-task, and eye-tracking. Nowadays, eye-tracking is one of the popular ways of measurement in the field. Eye-tracking involves observing participants' directions of eye gazes over time in response to a particular task. There is a meta-analysis on eye-tracking relating to speech recognition. The studies reviewed by Van Engen and McLaughlin (2018) found that general eye-tracking studies show listening ability decreases in noisy conditions, with accented speech, or when participants have hearing implants.

An alternative to eye-tracking software that can be implemented without specialized equipment is mouse-tracking. Mouse-tracking software tracks mouse movement on the screen and analyzes the mouse trajectory during computerized experiments. Both of these methods measure the participants' attention to the speech and their processing time. Just as the time course and pattern of eye movements can reflect the listening effort, the trajectory of the mouse cursor and response time can similarly reflect it. Laura Barca and Giovanni Pezzulo (2012) used a mouse-tracking method to measure listening effort. Using a "lexicality effect," which shows better performance with lexical sound than with non-lexical sound, they proved that a mouse-tracking method yielded the same results as eye-tracking in measuring listening effort (Barca, L., Pezzulo, G., 2012). Our team used the mouse-tracking method to measure listening effort in our experiment.

After analyzing the previous use of mouse-tracking, and the correlation between the native language and sentence context, our team hypothesized that English speakers from the United States would be able to benefit from constrained sentences. In contrast, South Korean participants, who are not continually exposed to American-accented English, would benefit less.

Method

Subjects

Our study used a convenience sample of 60 people across three different University Campuses: The University of Utah (Utah, USA), The University of Utah Asia Campus (Incheon, South Korea), and Ghent University Global Campus (Incheon, South Korea). This participant number was our final after adjusting for incomplete and duplicated responses. We collected 44 participants from the United States and 16 participants from the two South Korea Campuses. We chose to separate the groups by geographical location rather than their first language for our study. This decision was made by examining what accented English the participants were most exposed to on their respective campuses.

The Utah, USA population's birth years ranged between 1974 and 2003, with a mean year of 1999 (SD: 6.041). 68.2% (30) identified as female and 31.8% (14) identified as male. Within this population, 95.5% (42) people rated themselves as having no Korean proficiency, while 4.5% (2) people rated themselves as having some Korean proficiency. 97.7% (43) people reported having 0-1 year of exposure to the Korean language, and 2.3% (1) person reported having three to five years of exposure to the Korean Language. 9.1% (4) reported being exposed daily, 4.5% (2) reported being exposed a few times a week, 4.5% (2) reported being exposed a few times a month, 20.5% (9) reported being exposed a few times a year, and 61.4% (27) reported never having been exposed. 100% (44) of participants have lived or studied for 0-1 year in a Korean-speaking country.

From the Incheon, South Korea population, the birth years ranged between 1996 and 2002, with a mean birth year of 1999 (SD: 1.84). 66.7% (10) participants identified as female, and 33.3 (5) identified as male. Within this population, 6.7%(1) participants self-reported being fluent in English, 13.3% (2) reported having high proficiency in English, 33.3% (5) reported being moderately proficient in English, and 6.7% (1) reported having some proficiency in English. 26.7% (4) participants reported being exposed to English daily, 26.7% (4) reported being exposed a few times a week, and 6.7% (1) reported being exposed a few times a month. 20.0% (3) reported having lived or studied in an English speaking country for 0-1 years, 26.7% (4) have lived in an English speaking country for 1-3 years, 6.7% (1) lived in an English speaking country for 3-5 years, and 6.7% (1) lived in an English speaking country for more than five years.

Equipment

All data for this study was collected through the Qualtrics Survey platform, using four counterbalanced questionnaires. The response time data and pre-questionnaire responses were recorded through an Open-Source JavaScript code embedded into the survey (Mathur & Reichling, 2019). Each participant used their personal computer and speakers/headphones to complete their responses.

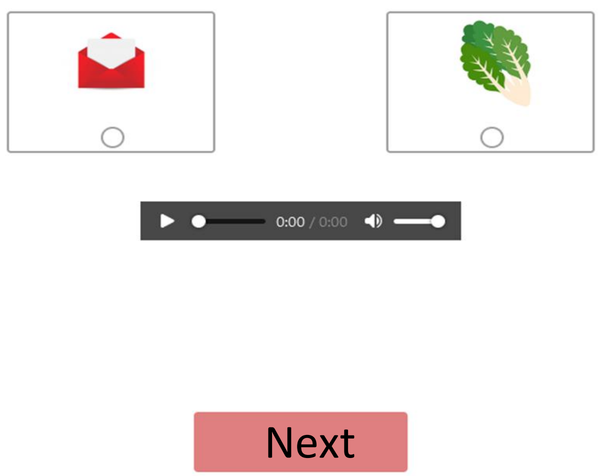
Because of the JavaScript program used, we could not use the random order function to balance our survey. Instead, we created four surveys with fixed random orders (determined by a random number generator) and assigned each participant to one of the four surveys based on their birth month.

Procedures

To adhere to Covid-19 guidelines and collect from a global sample, the data for this experiment was performed online. Each subject answered a 13-question pre-questionnaire that asked them to self-report their first language, their language proficiency, any known hearing or auditory processing issues, and the device specifications that they were using. Each participant was then presented with four sample questions (two for the first section and two for the second section) that gave the participant a chance to familiarize themselves with the software.

For the first task, each participant completed 30 trials, in which they were shown two images that depict everyday objects on their screen. Each pair of objects began with identical phonemes and depicted similar sounding words (EX: Money vs. Monkey) (See Appendix A for the complete target word list). At the same time, an audio clip of a sentence played in American accented English that included one of the two items shown. These audio clips were randomized between unconstrained sentences with no contextual clues (EX: "Point to the bucket") or a constrained sentence that included contextual clues (EX: "The man filled the bucket"). (See Appendix B for the complete unconstrained and constrained sentence list). Each participant began with their cursor on the "next button" (see Figure 1) while listening to the auto-played audio clip, then moved their cursor to one of the two Target word images and selected it. The response time began recording when the audio clip began to play and stopped when they selected one of the two images.

Figure 1: The User Interface of the Qualtrics Survey



Note: The JavaScript embedded code allowed us to program each box equidistant apart from one another to avoid confounds in the data.

After completing the first task, the participants performed a second, similar task. The results of this secondary task are outside the scope and discussion of this paper. The total duration of the experiment was ~30 minutes.

Results

The response times (RTs) collected in the study were analyzed using linear mixed-effects (LME) models in the R analysis software. The model that was the best fit was:

*Rt ~ condition * group * trial NR + (1 | ppNR) + (1 | stimNR)*

The RTs were modeled as a function of condition (const/uncon), group (US/SK), and presentation order (trial 1…30), with random intercepts for each participant and each audio recording.

The intercept represents the reference conditions: United States participants, constrained sentences, trial NR0, and each main effect or interaction can be added to this value to obtain an estimate for the other condition/group/trial (See Table 1).

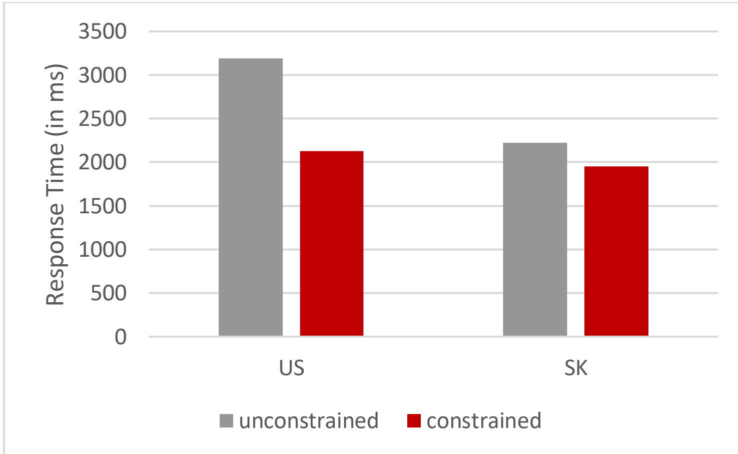
Table 1: Results of the Linear Mixed-Effect (LME) Models

	Estimate	Std. error	df	p-value	sig
(Intercept)	2124.348	237.579	199.624	2.61E-16	***
Cond-Uncon	1065.567	212.692	1665.129	6.02E-07	***
groupSK	-169.748	438.82	186.742	0.6993	
trialNR	-19.536	9.182	1559.6	0.0335	*
cond-Uncond:groupSK	-963.204	425.489	1670.709	0.0237	*
cond-Uncond:trialNR	-52.91	12.029	1671.296	1.16E-05	***
groupSK:trialNR	-7.124	17.289	1672.712	0.6803	
cond-Uncond:SK:trialNR	53.442	23.963	1673.618	0.0259	*

Note: *The result of our LME analysis indicated a training effect (trialNR) of -19.536 milliseconds per question that affected our mean response time. Our analysis also revealed an interaction effect between unconditioned sentences, South Korean participants, and the training effect.*

In our initial analysis of mean response times, we saw an average response time from United States-based participants of ~3200 milliseconds for unconstrained sentences versus a ~2200 millisecond response time from South Korean participants. For constrained sentences, United States participants had an average response time of ~2100 milliseconds, while South Korean participants averaged ~2000 milliseconds (see Figure 2).

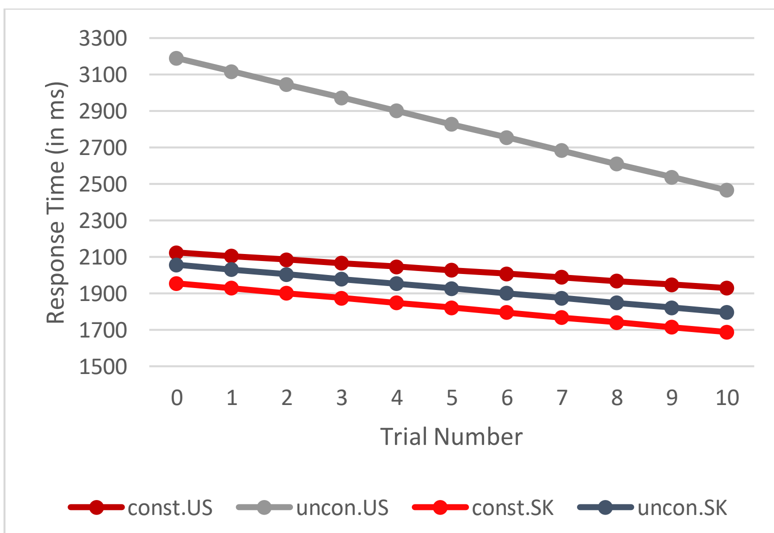
Figure 2: Average Response Time (Before Training Effect)



Note: This table represents the mean response time for each participant group before factoring in the training effect, or TrialNR of -19.536.

After factoring in the training effect of -19.536 milliseconds per response, we saw a significant decrease in response time for unconstrained response time from United States participants. However, there was no substantial change to constrained response time from United States participants or any responses from South Korean participants (see Figure 3).

Figure 3: Training Effect



Note: After factoring in the training effect, there is a significant change in unconstrained United States responses but no significant change from any other data group.

Discussion

The data collected in our study significantly supported our hypothesis that United States participants who were continually exposed to American-accented English could benefit from constrained sentences, whereas South Korean participants who were not continually exposed did not.

As exemplified by our graphs, United States participants' RT was significantly faster than unconstrained sentences for constrained sentences. This data suggests that they can benefit from the context in sentences while performing the task. South Korean participants did not show the same benefit. Additionally, United States participants showed a strong learning effect for unconstrained sentences over time. Again, South Korean participants did not show the same results. This information could be interpreted that South Korean participants benefitted less from sentence context in unfamiliar US accented speech.

This result is directly supported by the previous literature of Oosthuizen et al. (2019). They noted that response times are more closely affected by what language a person is exposed to rather than their native language, which supports our results based on the participant's geographical location. Van and McLaughlin's 2018 research also directly supported our results that accented language and unfamiliar accents directly affect processing time and cognitive load required to process auditory speech.

This information is significant in an international educational setting by providing valuable information about how students learn in a non-native language and the differences in comprehension efforts compared to their native-speaking peers. By understanding the listening effort needed, educators can format their approaches to learning to accommodate this difference and provide an equitable education for all students. As a global partner University with campuses in both the United States and South Korea, this is especially significant to educators at the University of Utah and is crucial knowledge for global education and exchange.

Limitations and Future Directions

Because the study was completed through a convenience sample, and there was a significant disparity between the number of United States participants and South Korean Participants, this study was relatively limited in scope. This limitation also directly affects our external validity. For more comprehensive future studies, participation from a broader range of participants would be necessary. This change would mitigate any confounds further and factor them into future results.

Internal validity could have also been threatened in this study, as all participants took the study virtually, not in a controlled environment. This issue could affect the attention of the participant, any distractions in their respective environments, and the amount of working memory they dedicated to the survey. This, in turn, could affect how quickly they could cognitively process the information and respond.

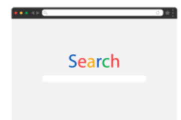
Based on the nature of the results, it is also possible that the data reflects a floor effect for South Korean participants. Their constrained and unconstrained response times closely mirrored United States participants' constrained (improved) response times. This phenomenon suggests there may have been room for improvement in their response times, but they were limited by the software and nature of the survey. Based on the scope of data collected, we cannot isolate one specific cause of this phenomenon.

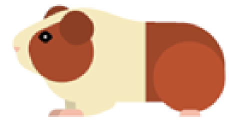
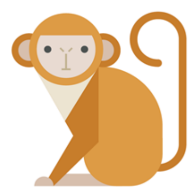
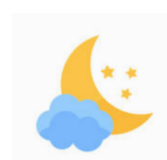
To correct the scope of both internal and external validity in future research, identical controlled environments should be used for both South Korean and United States studies. This will decrease varying confounds and ensure that each participant has the same level of focus and audio quality.

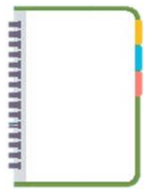
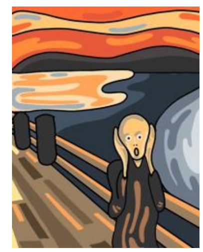
Language and the ability to process it are vital parts of our everyday lives. Previous research on how we process unfamiliar language has convergently shown that we require more attention to comprehend versus the few resources we use to process familiar language. Future research into this field could improve linguistic skills and provide more advanced educational opportunities for international students studying in a foreign language.

Appendix A

A Complete List of Target Words and Images Used in our Experiment

Target	Image (illustrations)	Target	Image (illustrations)
Baby		Baseball	
Brown		Browser	
Bucket		Buckle	
Butter		Button	
Candle		Cannon	
Cat		Cap	

Coffee		Coffin	
Hammer		Hamster	
Letter		Lettuce	
Medal		Medicine	
Money		Monkey	
Knife		Night	
Plane		Plate	

Planner		Planet	
Screen		Scream	

Note: All photos were acquired through an iStock educational use license.

Appendix B

A Complete List of Unconstrained and Constrained Sentences Used

Target Word	**Constrained Sentence**	**Unconstrained Sentence**
Baby	The man fed the baby.	Point to the baby.
Baseball	I hit the baseball.	Point to the baseball.
Brown	The leaves turned brown.	Point to the brown.
Browser	The man opened the browser.	Point to the browser.
Bucket	The man filled the bucket.	Point to the bucket.
Buckle	He clipped his buckle.	Point to the buckle.
Butter	I spread the butter.	Point to the butter.
Button	The girl fastened her button.	Point to the button.
Candle	The man lit the candle.	Point to the candle.
Cannon	She fired the cannon.	Point to the cannon.
Cat	She pet the cat.	Point to the cat.
Cap	He wore the cap.	Point to the cap.
Coffee	I drank coffee.	Point to the coffee.
Coffin	The woman buried the coffin.	Point to the coffin.
Hammer	He banged the hammer.	Point to the hammer.

Hamster	My sister pet the hamster	Point to the hamster.
Letter	The woman wrote the letter.	Point to the letter.
Lettuce	The man ate the lettuce.	Point to the lettuce.
Medal	My sister wore the medal.	Point to the medal.
Medicine	The man swallowed the medicine.	Point to the medicine.
Money	The man paid the money.	Point to the money.
Monkey	My brother fed the monkey.	Point to the monkey.
Knife	The man sharpened the knife.	Point to the knife.
Night	She slept at night.	Point to the night.
Plane	He flew the plane.	Point to the plane.
Plate	The woman washed the plate.	Point to the plate.
Planner	The woman wrote in the planner.	Point to the planner.
Planet	NASA discovered the new planet.	Point to the planet.
Screen	The man watched the screen.	Point to the screen.
Scream	The man heard the scream.	Point to the scream.

References

Barca, L., Pezzulo, G. (2012). Unfolding Visual Lexical Decision in Time. PLOS ONE 7(4):e35932. https://doi.org/10.1371/journal.pone.0035932

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models : a tutorial. Journal of Cognition, 1(1), 1–20. doi:10.5334/joc.10

Huettig, F., & Janse, E. (2015). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. Language, Cognition and Neuroscience, 31(1), 80-93. doi:10.1080/23273798.2015.1047459

Mathur, M.B., Reichling, D.B. Open-source software for mouse-tracking in Qualtrics to measure category competition. Behav Res 51, 1987–1997 (2019). https://doi.org/10.3758/s13428-019-01258-6

Nitsan G, Wingfield A, Lavie L, Ben-David BM. Differences in Working Memory Capacity Affect Online Spoken Word Recognition: Evidence From Eye Movements. Trends Hear. 2019 Jan-Dec;23:2331216519839624. doi: 10.1177/2331216519839624. PMID: 31010398; PMCID: PMC6480998.

Oosthuizen, I., Picou, E., Pottas, L., Myburgh, H., & Swanepoel, D. (2019, December 01).

Listening Effort in Native and Non-native English-Speaking Children Using Low Linguistic Single- and Dual-Task Paradigms. Retrieved October 19, 2020, from https://pubs.asha.org/doi/abs/10.1044/2020_JSLHR-19-00330

Pals, Carina & Sarampalis, Anastasios & Rijn, Hedderik & Başkent, Deniz. (2015). Validation of a simple response-time measure of listening effort. The Journal of the Acoustical Society of America. 138. EL187-EL192. 10.1121/1.4929614.

Peng, Z. E., & Wang, L. M. (2019). Listening Effort by Native and Non-native Listeners Due to Noise, Reverberation, and Talker Foreign Accent During English Speech Perception. Journal of Speech, Language, and Hearing Research, 62(4), 1068-1081. doi:10.1044/2018_jslhr-h-17-0423

Rönnberg, J., Stenfelt, S., & Lyxell, B. (2013). The Ease of Language Understanding (ELU) model: Theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience. 7.* https://doi.org/10.3389/fnsys.2013.00031

Van Engen Kristin J., Peelle Jonathan E. (2014). Listening effort and accented speech, Frontiers in Human Neuroscience. 8. https://doi.org/10.3389/fnhum.2014.00577

Van Engen, K., & McLaughlin, D. (2018). Eyes and ears: Using eye tracking and pupillometry to understand challenges to speech recognition. Hearing Research. 369. 10.1016/j.heares.2018.04.013.

Wagner, A., Pals, C., de Blecourt, C., Sarampalis, A., & Başkent, D. (2016). "Does signal degradation affect top-down processing of speech?" In P. van Dijk, D. Başkent, E. Gaudrain, E.

de Kleine, A. Wagner, & C. Lanting (Eds.), Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing. (Vol. 894, pp. 297-306).

Xie, X. & Fowler, C. A. (2013). Listening with a foreign-accent: The interlanguage speech intelligibility benefit in Mandarin speakers of English. Journal of Phonetics, 41, 369-378.doi:10.1016/j.wocn.2013.06.003