

**DIFFERENT REGIONAL VARIANTS OF ENGLISH  
AND THEIR INTERACTION WITH NATIVE  
PERCEPTION**

**Author:**

*Caleb Baker*

**Supervisor:**

Suki Yiu

**Second reader:**

Paul Boersma

**June 2021**



UNIVERSITEIT VAN AMSTERDAM

## **Acknowledgments**

Thanks go out to Suki Yiu for helping advise me over the entirety of my time spent writing, Paul Boersma for being the second reader, Rorick Terlou for moral encouragement as well as advising. I would like to thank the participants, who were all calm and cooperative, with some being especially interested in the research. I would not have been able to finish this without the help of every single one of these people, and for that they have my sincerest gratitude.

## **Abstract**

The goal of this thesis is to determine if there is a tangible link between the nativeness of an accent that was spoken and the reaction and accuracy of participants when answering questions about those spoken stimuli. The chosen area for researching this was emphasis, as it has a very distinct way of making itself known, and it could be evaluated which words in the experiment stimuli were emphasized as to give a correct and incorrect answer to the emphasis-based questions for the accuracy pieces. The experiment was run in Qualtrics, and then the results were run through an LMER and GLMER test in R to test the reaction time and accuracy respectively of each accent speaking group (Southern North American, Mid-Western North American, Australian and London English). The results showed that only speakers of the Mid-Western North American regional dialect reported any significant difference between questions in their native accent and in a non-native accent, while other accent groups were not found to be significant in relation to the nativeness of the accents that they were presented with. An explanation that seems likely is the prevalence of Mid-Western North American accent media in today's world. This can be seen as far between as the educational videos schools use to the preferred accent of television hosts. The prevalence of a believed Standard American English (SAE) can also be attributed to this behaviour, as the majority of people believe that the Mid-Western North American accent is this SAE.

## Table of contents

<b>Title</b>	<b>0</b>
<b>Acknowledgments</b>	<b>1</b>
<b>Abstract</b>	<b>2</b>
<b>Table of Contents</b>	<b>3</b>
<b>1. Introduction</b>	<b>4</b>
<b>2. Literature Review</b>	<b>5</b>
2.1 Emphasis	5
2.2 Emphasis in Southern North American and Mid-Western North American accents	6
2.3 Emphasis in Australian and English accents	6
2.4 Research question and predictions	7
<b>3. Methodology</b>	<b>7</b>
3.1 Participants	7
3.2 Stimuli of the four varieties of English	8
3.2.1 Stimuli explanation	8
3.2.2 Southern North American: Clint Eastwood (acting accent)	9
3.2.3 Mid-Western Northern American: Kevin Costner (acting accent)	9
3.2.4 Australian: Paul Hogan (authentic accent)	9
3.2.5 London English: Benedict Cumberbatch and Martin Freeman (Authentic)	10
3.3 Procedure	10
3.4 Data analysis	10
3.4.1 Reaction time graph (scatter plot)	11
3.4.2 Accuracy graph (stacked bar plot)	12
<b>4. Results</b>	<b>12</b>
4.1 Southern North American	13
4.1.1 Reaction time	13
4.1.2 Accuracy	14
4.2 Mid-Western North American	15
4.2.1 Reaction time	15
4.2.2 Accuracy	16
4.3 Australian	17
4.3.1 Reaction time	17
4.3.2 Accuracy	18
4.4 London English	19
4.4.1 Reaction time	19
4.4.2 Accuracy	20
4.5 Results summary	20
<b>5. Discussion</b>	<b>21</b>
<b>6. Conclusion</b>	<b>22</b>
<b>References</b>	<b>23</b>
<b>Appendices</b>	<b>24</b>

## **1. Introduction**

In today's ever integrated world, it is not uncommon to come into contact with your own native language being spoken in a way that is dissimilar to your own. It can be difficult to understand someone else's accent and subsequently it is assumed that because they have a different accent it will be more difficult to react to that accent in comparison to your own. This question is the crux of this thesis, and is the main point that will be discussed. How effective are nativeness and familiar accents in creating faster and more accurate reactions? With this goal in mind, the first step is to find a medium in which to test the differences in perception of a person's native language.

Emphasis is this medium, and is also a matter of great importance for perceiving language in general. Even if you personally do not speak a language, you might still be able to comprehend which parts are important, and which are not. This has to do with a person's understanding of universal signs of emphasis, and is interesting when looked at through the lens of different accents. If the recognition time of accents in a language a person speaks natively are slower when encountering an accent that they do not speak with, is their recognition of the emphasized parts of a phrase also similarly hampered, or do the internal rules of emphasis override the slower perception speed? This could be the key to pinning down the importance of emphasis in cross-accent conversations, and has implications for situations in which speed is of the necessity.

Baker (2020) ran an experiment about the perception of final voicing and it was the spark that piqued my interest in perception studies. Along with my already vested interest in accents, it seemed prudent to merge the two ideas. This ultimately led to me wanting to determine how easily native english speakers from different regions are able to perceive emphasis from the same and different accents. This will be accomplished by setting up a reaction time and accuracy experiment based around emphasis, and then a comparison of the results in R to test if there is a connection between the nativeness of an accent to a participant and their reaction time and accuracy.

## **2. Literary review**

### **2.1 Emphasis**

Sophie Herment and Daniel Hirst (2002) wrote about how emphasis is discerned in English. This study was conducted on native English speakers, and had participants identifying what was an emphasized part of a presented phrase . The results of the study were separated into four sections based on how the emphasis was created (pauses, duration, fundamental frequency, and semantics). Of the four categories, it was seen that, in the perception of emphasis, the fundamental frequency and semantics were paramount in distinguishing emphasis. It should also be noted that while duration is not a necessity for perceiving emphasis, it still does have an effect on that perception according to the study. Based on these results, the thesis would be an excellent choice for creating the stimuli of this thesis's experiment as it specifies how to find the emphasis in a phrase for the accuracy portion. Hermit and Hirst (2002) also accomplish what this paper is trying to discover in relation to emphasis perception, without overlapping concept-wise as the authors' main focus was only on emphasis in English while mine is based around using emphasis to determine how nativeness of an accent can affect the response time and accuracy of different native English regional dialect/accents speakers. The main objective of using this paper can be accomplished easily, but it will be difficult to not step on the authors' proverbial toes when conducting the actual experiment.

Since accent will be focused on as the way emphasis is used in this study, a document on it as well helped flesh out my idea of how to incorporate emphasis. Sarah Thomas and James Scobbie (2015) wrote about the first language acquisition of children of Scottish descent that were adopted into English families after their first year. The paper mostly focuses on how the change from speaking with a Scottish accent at home and being educated with an English accent at school can cause a rift in perception. Mainly, it was found that these mixed-accent children were able to better understand the accent presented to them if they were in a location associated with that accent (i.e. Scottish at home and English at school).

## **2.2 Emphasis in Southern North American and Mid-Western North American accents**

Amalia Arvaniti and Gina Garding (2007) discovered the basis for how the accents of Southern North American and Mid-Western North American show emphasis in spoken language in their paper *Dialectal variation in the rising accents of American English*. The results of the experiment performed by Arvaniti and Garding display that the accents are similar in creating spoken emphasis, but differ in the rising tones in regard to how they transition from a H\* to a L+H\* tone. It was seen that there was not enough evidence for the Mid-Western North Americans distinguishing between the two tones, but kept a steady L+H\* where there was a difference, whereas the Southern North American speakers did distinguish the difference in emphasized speech. This shows that while the display of emphasis is mostly the same, there are exceptions and differences that might impede perception in this experiment.

## **2.3 Emphasis in Australian and English accents**

Qin Yan and Saeed Vaseghi (2003) illustrate the differences in formants, and their effect on the accent perception between the three accents of British, Australian, and American English. The findings of their study state that the first formants were higher in British and American English dialects in comparison to the Australian accents, and that the Australian accents had a higher second formant than the American and British English accents. This difference in formant pitch shows that there are differences in how words are pronounced, and in turn which vowels carry emphasis, within the three language accents, but especially between the British English and Australian variants. Also, it is seen that the durations of the emphasis are closer between the British English and Australian accents than with the North American accents and vice versa between the Southern North American and Mid-Western North American accents. Therefore, this paper is used as the theoretical justification for the differences in emphasis between the two accents of British English and Australian.

## **2.4 Research question and predictions**

When taking into account the research question of this study, how easily are native English speakers from different regional dialects able to perceive emphasis from native as opposed to non-native accents, there are a few predictions that can be made as to the outcome of the experiment and its findings. Mainly, it is predicted that a participant's response speed and response accuracy in relation to the emphasized words will be faster and higher respectively to how close the regional accent is to their own geographically. This would mean, based on a geographical logic stance, that the results from the Southern North American and Mid-Western North American speakers should be closer than the London English and Australian native speakers and vice versa. Also, if this theory is considered reasonable, then the Australian accent results should be closer to the American accent varieties than with London English. Therefore, based on this information, it is assumed that a participant will be the fastest and the most correct when they are presented with an audio clip that uses their own personal regional native English accent, and close to this native speed and accuracy when presented with an accent close to their native one geographically.

## **3. Methodology**

### **3.1 Participants**

The participants were aged 18 and up. These subjects spoke English natively, had no mixed accents (i.e. they speak in the same accent they were born into), and had not had any auditory or visual disabilities. The accents studied were five Southern North American speakers, seven Mid-Western North American speakers, fifteen London English speakers, and eleven Australian speakers.

Deutskens et al. (2004) examines how taking a survey online could change the response rates of the participants. By replicating other offline shopping surveys, the authors created an experiment online and conducted it in the hopes that it would shed light onto the potential increase in response rates of participants. This was conducted over email and the results showed a dramatic increase in responses from participants in comparison to the survey's mail-in counterparts. This study showed the relevance of internet surveys being a legitimate medium to



conduct an experiment because of the high participant response rate. It is also the basis I took for deciding to distribute the experiment over email to my participants.

### **3.2 Stimuli of four varieties of English**

The stimuli for the experiment consists of twenty-four 5 second audio clips from the movies *The Good, the Bad, and the Ugly* (1967), *Man of Steel* (2013), *Crocodile Dundee* (1986), *Crocodile Dundee II* (1988), as well as the first season of the BBC television series *Sherlock* (2010). The actors and accents that were taken from these pieces of media are as follows respectively; Clint Eastwood (Southern North American), Kevin Costner (Mid-Western North American), Paul Hogan (Australian), Benedict Cumberbatch and Martin Freeman (London English). These clips are audio only, twelve of which are used because of the basis of this thesis as they employ the use of emphasis to convey an idea to the listener. The remaining twelve audio clips will be used as filler, clips without emphasis being expressly employed, so that participants do not become primed towards emphasis. The emphasis was decided by the head researcher based on the theoretical background covered above, and through the use of formant evaluation in the program of Praat. These audio clips are from five speakers, who employ different accents, whether through acting or by using their natural accent. This decision was made based on the piece *Ontology of Fake: Discerning the Philippine Elite* by Angela Reyes which states that “an ontology of fake presupposes the existence of types whose tokens aspire to pass for tokens of differentiated types through double-voiced emblematic displays”. Essentially, this means that the differentiation of acted and authentic is superficial as long as the display of the fake is taken as authentic. A complete list of these speakers can be seen below with a description of their background.

#### **3.2.1 Stimuli explanation**

This experiment consisted of twenty-four duos, all of which will only be used once each. These duos consist of a slide for the stimuli which are audio clips in a random native English regional accent of the four stated before and a question slide asking the participant to choose the emphasized word of the stimuli phrase from four potential choices. These twenty-four duos will be randomized for each participant, and the slides will have a maximum time limit of twenty seconds added to them to decrease the likelihood of abnormally long times to answer (such as

leaving to go to the toilet etc.). Twelve of the twenty-four questions asked to the participant are filler questions, meaning they are there to limit the priming of the emphasis in the target questions. The filler questions do not have any right answers, but the target questions do. A full list of the audio clips and their respective links can be found in the appendices.

### **3.2.2 Southern North American: Clint Eastwood (acting accent)**

Clint Eastwood's clips from the movie *The Good, the Bad, and the Ugly* (1967) are the source for the Southern North American accent. Clint Eastwood himself was born and raised in San Francisco, California, and as such, does not natively speak with a Southern accent. Despite this, I still preferred to choose him as his acting of the accent was indiscernible in the clips used from the actual accent, and as the perception study is based on the participants' reactions to the accent, it need not be a genuine accent as long as it is close enough for native speakers to find it indiscernible from a native one. Six audio clips were used from Clint Eastwood in the movie.

### **3.2.3 Mid-Western Northern American: Kevin Costner (acting accent)**

Kevin Costner's voice from the movie *Man of Steel* (2013) was used as the source of the Mid-Western Northern American accent. Similar to Clint Eastwood before, Costner was originally born in Lynwood, California, and therefore is acting the Mid-Western accent in the movie. It should be noted however, that Costner's upbringing was closer to the Mid-Western North American accent than Eastwood's to the Southern North American Accent. For parallel reasons to Eastwood's; it was determined to have him represent the accent that he was acting. Six clips were used from Kevin Costner from the movie.

### **3.2.4 Australian: Paul Hogan (authentic accent)**

Paul Hogan's voice from his appearance in the films *Crocodile Dundee* (1986) and *Crocodile Dundee II* (1988) were used for the sound clips with an Australian accent. Hogan was born and raised in Sydney, Australia, and therefore speaks the Australian accent natively. He was chosen because of his easily distinguishable Australian accent to any native speaker of English. Six audio clips were used from Paul Hogan across the two movies, two from the first, and four from the second.

### 3.2.5 London English: Benedict Cumberbatch and Martin Freeman (Authentic accents)

Benedict Cumberbatch and Martin Freeman were the choices to represent the English accent from their roles in the first season of the BBC television series *Sherlock* (2010). They were born and raised in London, England and Hampshire, England respectively. This means that their accents are genuine and not acted. They were chosen similarly to Paul Hogan for their distinct voices and easy to recognize accents to native English speakers. Two audio clips are used from Benedict Cumberbatch and four from Martin Freeman.

### **3.3 Procedure**

Participants were presented with a Qualtrics experiment. Before the experiment begins, the subjects were given an information brochure detailing a brief and general overview of the experiment's purpose, as well as a consent form in which they needed to select a box stating that they agree to said consent form before the participant was allowed to start the experiment. The study begins with a slide informing the participants to listen to the stimuli and to pick the emphasized part of the stimuli on the next slide as quickly as they can, as well as the fact that the survey will take about twenty to twenty-five minutes.

The results of the time it took for participants to answer the questions as well as the accuracy of the participants in choosing the right answers was collected into an Excel spreadsheet on an encrypted usb drive. This was to ensure that the privacy of the participants is maintained, even though the names of the participants were not collected at all. The reaction time of the participants in Qualtrics is reliable as the audio clips are uniform in length and each question the participants are presented with has a maximum time limit to ensure that they are answered in as fast a time as possible. This sense of urgency is to better facilitate the idea of participants reacting as fast as they can to achieve an accurate reading of their actual reaction time ability.

### **3.4 Data analysis**

The data was run through two separate models for every regional dialect group; four groups in total. The models that were used were a Linear Mixed Effect Model (LMER) and a Generalized Linear Mixed-Effects Model (GLMER). These tests were used because they could account for

the independent variable of the regional accent a participant had, the dependent variable of that participant's reaction time and accuracy, and finally for the random aspect of which participant was being analysed. Then, the results of each group were plotted on a scatter plot graph to analyze reaction time and a stacked box graph to analyze accuracy. These graphs were chosen for their simplicity as well as their ability to convey the information of relevance accurately. The scatter plot was used for the visual element of the reaction time data because it supported more than two options on the x-axis (stimuli regional accent) as well as multiple options for the y-axis (reaction time with stamps showing the minimum number up to the maximum reaction time). The stacked box plot was preferable for accuracy because of its binary method of displaying the x-axis (Correct or Incorrect) as well as the organized nature of the different stimuli regional accents that the participant answered correctly or not on. To create these graphs, the results were placed into a Microsoft Excel sheet, and then the data was manually arranged so that the file could be imported as a .csv file into R (R core team, 2021). The model creations, as well as the graph creations for each of the regional accent groups was done within the R (R core team, 2021) program. The full code for the tests and graphs can be found in the appendices.

### **3.4.1 Reaction time graph (scatter plot)**

A scatter plot was the graph of choice for displaying the reaction times of the participants. Four scatter plots were made, one for each native accent group of participants, and then their reaction times were inserted into the graph. The reaction time, which peaks at twenty seconds as that is the maximum amount of time allotted for the participants per reaction, makes up the y-axis of the graph, while the x-axis is split into two variables, native and non-native (seen in the variable AccentN) referring to the nativeness of the accent that the reaction was in relation to. A blue mean value point is also inserted into the graph, to show what the average reaction time was in relation to the values of the x-axis. The legend displays the initials of the native accent group (which is symbolized by the variable Languagelistener) that the graph is showing which consist of Southern North American (SNA), Mid-Western North American (MWNA), Australian (AUS) and London English (E). A secondary scatter plot shows the inter-accent comparison of the chosen group in relation to all the stimuli accents (shown through the variable Languagespeaker on the x-axis). This graph was chosen because it can display many different points at once,

which facilitates the analysis of the answers, as well as giving the ability to split the reactions between the nativeness of the accent the answer was for and the participants' reaction speed.

### **3.4.2 Accuracy graph (stacked box plot)**

A stacked box plot was used to display the data of the accuracy of the participants. Similar to the reaction time analysis, four graphs were created, one for each native accent group of participants. The y-axis displays the initials of the native accent group that the graph is showing, with the same initials as the legend in the scatter plot. The x-axis is separated into four columns, one for each accent stimuli that the group gave answers for (shown through the variable (Languagespeaker). Finally, the Legend shows which colour corresponds to correct and incorrect answers. This graph was chosen because of its neat design and the ability to show all the accents' results in relation to the participant (Languagelistener) group, which is on the y-axis. It also gives the ability to easily see which accents a particular group found easier or harder to identify the emphasis in.

## **4. Results**

The predictions for the experiment result were that the participants would be more likely to recognize the emphasized word faster and more accurately when it was presented in their native regional English accent than with the other three regional accents. For the reaction time graph, the minimum reaction time was the lowest reaction time that was seen within a regional dialect listener group and the maximum was twenty seconds, the longest time the participants had to answer the question before it auto-completed and they were ushered to the next question in the experiment. The accuracy graph shows the incorrect and correct answers of a regional dialect group in relation to all other dialects looked at.

## 4.1 Southern North American

### 4.1.1 Reaction Time

Figure 1: Nativeness

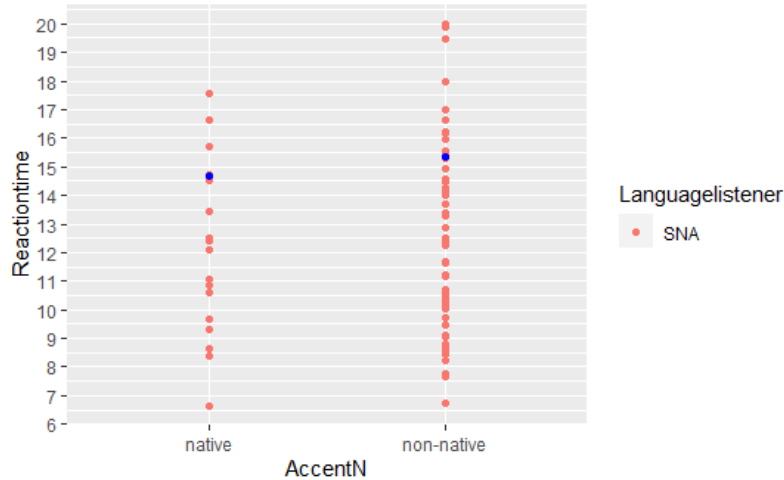


Figure 2: Inter-accent Comparison

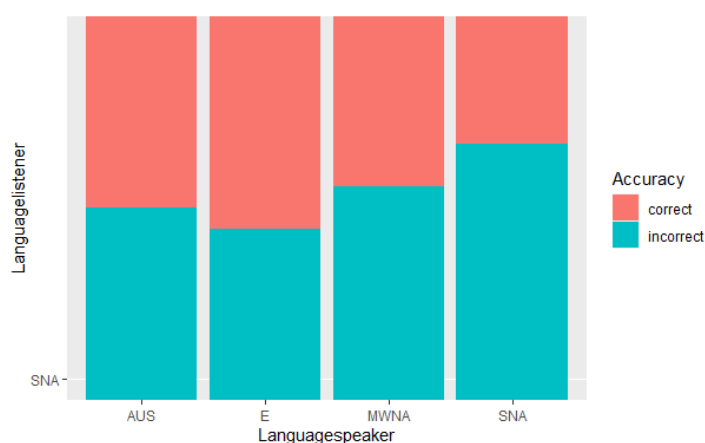


The Southern North American regional dialect speakers, based on the scatter plot from Figure 1, display a possible inclination to faster reaction times in their native regional accent as displayed in the lower overall native mean reaction time shown as the blue points on the graph. However, the reaction speed of the Southern North American speaker participants was mainly inconclusive according to the visual data of the scatter plot in Figure 2. With the independent variable being the native accent of the participant, the dependent variable being the reaction time, the nativeness LMER test showed a t-value of 0.66 and a p-value of 0.51 with 95% confidence intervals of  $[-1.04s, +2.08s]$ . The point estimate was 0.52 and the standard deviation was 2.09. These results

report that there is not enough evidence of a correlation between the Southern North American accent speakers, native accent and reaction time.

#### 4.1.2 Accuracy

Figure 3: Accuracy



For accuracy, the Southern North American regional accent group showed no significance value in relation to nativeness according to the nativeness GLMER test that was conducted. The stacked bar graph above as well appears inconclusive, as there is not a significant enough distance between the heights to the correct and incorrect columns. It should be noted, however, that the Southern North American accent speaking participants were the most inaccurate when confronted with questions pertaining to their own native accent. For the nativeness GLMR test, the p-value produced was 0.22. The 95% confidence intervals were [-1.9,+0.40] with an estimate value of -0.69. These results indicate that there is not enough evidence of a correlation between nativeness of accent and accuracy in relation to the Southern North American participant group.

## 4.2 Mid-Western North American

### 4.2.1 Reaction Time

Figure 4: Nativeness

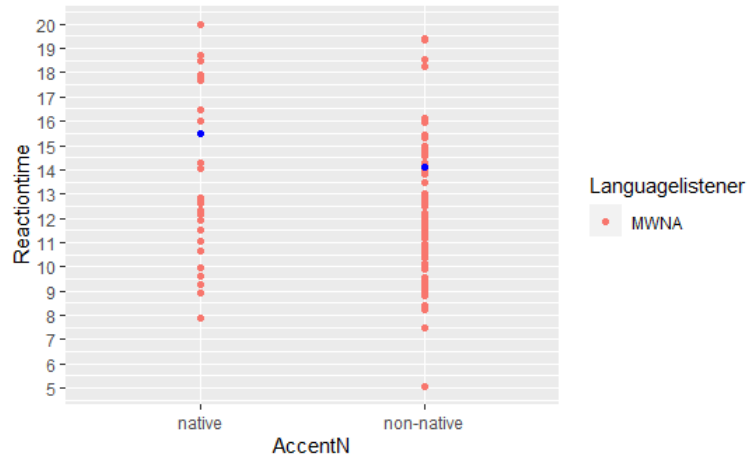


Figure 5: Inter-accent Comparison



The Mid-Western North American dialect speakers, according to Figure 4's mean values, show a faster mean reaction time overall when they were presented with non-native regional accents than with their own native accent. They didn't display a faster reaction time for any one regional accent that was tested besides an outlier within the Australian stimuli based on Figure 5's graph. However, the nativeness LMER test displays that there is a correlation between the native region accent speakers of Mid-Wester North American and their subsequent accent stimuli, showing a t-value of -2.0, a p-value of 0.040, 95% confidence intervals of [-2.861s, -0.078s], an estimate value of -1.5 and a standard deviation of 0.53. This means that the LMER test reports that there



is a correlation for native Mid-Western North American speakers' reaction time on determining emphasis when it comes to their native regional accent.

#### 4.2.2 Accuracy

Figure 6: Accuracy



Based on the figure above, it can be seen that there seems to be a higher number of correct answers than incorrect answers when the Mid-Western North American native regional accent speakers are presented with emphasis questions in their native accent. This observation from the graph is supported by the nativeness GLMER test that was conducted, which shows a p-value of 0.0034, 95% confidence intervals of [0.56,2.66] and an estimate value of 1.6. This means that the Mid-Western North American native accent speaking participants were indeed more accurate when posed with emphasis questions in their own native accent.

### 4.3 Australian

#### 4.3.1 Reaction Time

Figure 7: Nativeness

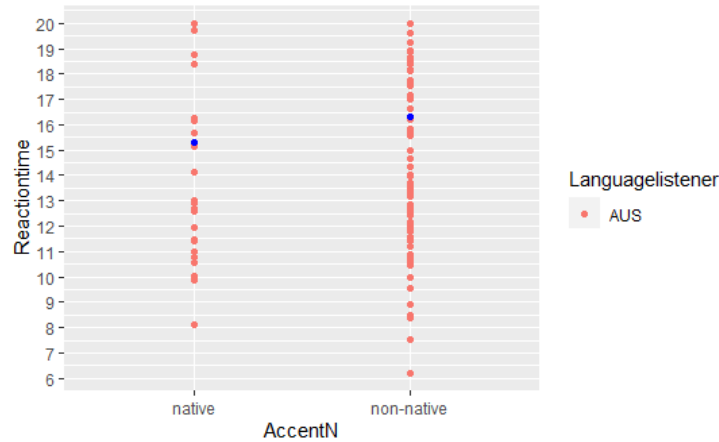


Figure 8: Inter-accent Comparison

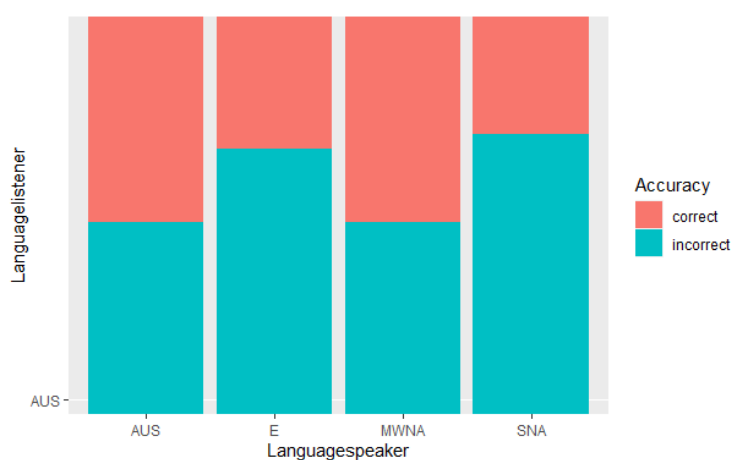


The scatter plot in Figure 7 for the native Australian regional accent results shows a slight proclivity between the participants' native accent and a faster reaction time when posed with discovering the emphasis of a phrase within that regional accent. However, Figure 8 shows that there might be faster overall emphasis recognition reaction time when this group was exposed to stimuli with a Southern North American regional accent. As for the nativeness LMER test for the Australian regional dialect group, it appears that there is not a significance in the reaction time in relation to the stimuli of their native regional accent. The t-value was 1.5, the p-value was 0.14, with the 95% confidence intervals  $[-0.30s, +2.08s]$ , an estimate value of 0.89 and a standard deviation of 2.07. These results report that there is not enough evidence of a correlation between

being a native Australian regional accent speaker and the speed at which one can distinguish emphasis from that same regional dialect.

### 4.3.2 Accuracy

Figure 9: Accuracy



The graph above of the accuracy of the native Australian regional accent participants shows that overall they answered with mostly incorrect emphasis options. A point of interest is that the accuracy of this group in relation to the Southern North American regional accent was that it was the most likely above the other three accents to be incorrect, especially in relation to its geographically closer accent, Mid-Western North American . The nativeness GLMER test that was conducted with this group's data revealed that overall there was not enough evidence of a correlation between nativeness of the Australian regional dialect and accuracy in understanding emphasis. This is shown by the p-value of 0.21, the 95% confidence intervals of [-0.32,+1.47] and the estimate value of 0.56.

## 4.4 London English

### 4.4.1 Reaction Time

Figure 10: Nativeness

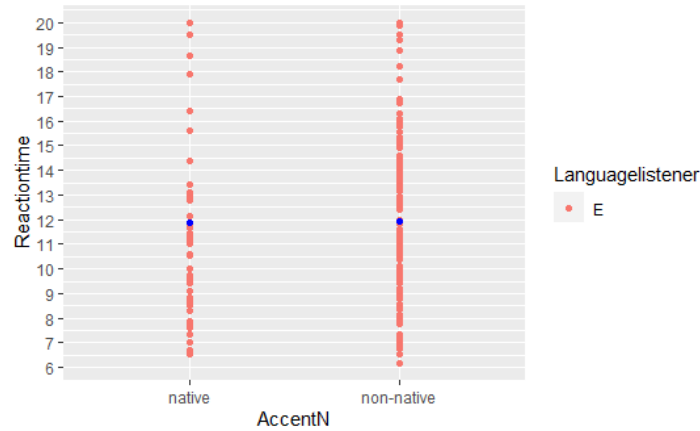


Figure 11: Inter-accent Comparison

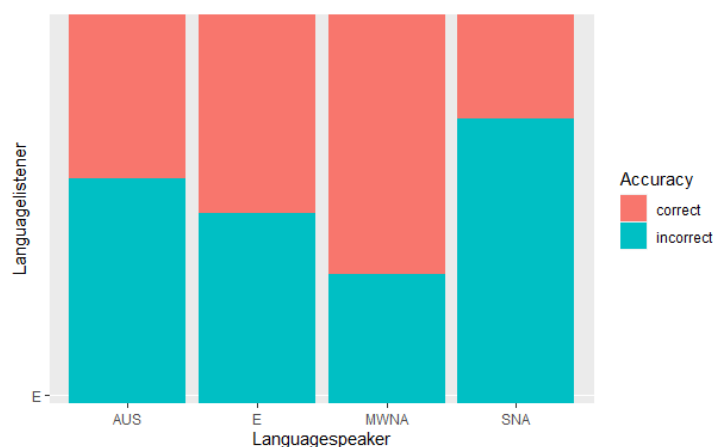


In Figure 10, it shows that the native speakers of the London English regional accent are overall barely faster in their own native accent than they are with the other non-native accents. The mean of the native reaction times was 11.875 seconds, and the mean for the non-native reaction times was 11.905 seconds, resulting in a difference of only 0.03 seconds. The native London English regional dialect group’s scatter plot in Figure 11 would suggest that they are the fastest out of the four groups overall when it comes to recognizing emphasis in the other regional dialects. The group also is shown to be drastically faster at identifying emphasis in the Australian regional dialect of English. Importantly, this proficiency in emphasis acquisition does not seem to dramatically extend to their own native regional dialect, at least based on the results of the

nativeness LMER test. This test showed a t-value of 1.93, a p-value of 0.054, 95% confidence intervals of  $[-0.014s, +2.001s]$ , an estimate value of 0.99 and a standard deviation of 1.3. These results show that there is not enough evidence of a correlation between the native London English speaking participants and the nativeness of the accent they were presented with in relation to their reaction time on emphasis based questions.

#### 4.4.2 Accuracy

Figure 12: Accuracy



The graph above is a visual representation of the London English regional accent speaker group's results. An aspect of note is that the propensity for the group to be correct on emphasis was highest when it was in the Mid-Western North American regional accent. The results of the nativeness GLMER test conducted state that the p-value was 0.48, the 95% confidence intervals were  $[-0.45, +0.96]$  and the estimated value was 0.25. This would indicate that there is not enough evidence of a correlation between nativeness of this accent and the accuracy of which this group of native accented participants answered the emphasis questions present in the experiment.

#### 4.5 Results summary

In summary, the results of the nativeness LMER and GLMER tests showed that the only accent participant group that displayed a correlation between their nativeness accent-wise relative to the accent of the stimuli and their reaction time and accuracy in answering the emphasis based questions were those within the Mid-Western North American group. This also conversely means

that the other three accent participant groups did not show a correlation in respect to their native accent, the accent of the stimuli, their reaction speed and their accuracy.

## 5. Discussion

There are a few explanations for the results witnessed from the experiment, but what I believe is the most crucial is the findings based around the Mid-Western North American accent. This regional dialect can be seen as having the most consistent results when it comes to accuracy and reaction time within all the participant accent groups, especially in the Southern North American and Mid-Western North American accent groups. Based on research, it is possible this is due to the fictional belief of Standard American English not only being propagated in the Mid-Western part of the United States of America, but much of the surrounding regions as well. As stated by Nancy Niedzielski (2002) in her chapter *Attitudes Toward Midwestern American English*, “Standard American English (SAE) has been erroneously attributed, even by some linguists, to the Midwest, New England and “the variety used by network news anchors”(Fromkin and Rodman 1983:87)”. She goes on to posit that the reason for the holding of this view by the American public is due to the low amount of overt stereotypes of the accent in relation to its neighboring accents.

This is the core of the reason for the importance of the accent within this study, particularly the propagation by the speakers of the accent as the default American accent, but also the fact that television recognizes this widely believed idea. As stated by Niedzielski, news anchors, reporters and other television branches want to appeal to as wide a range of audience as possible to gain viewership. This goal has led most of the television and media in general to be given to consumers in a Mid-Western North American accent. The more contact a person has with an accent, the more easily they will be able to understand it. Tracy Derwing and Murray Munro (2009) argue that over time, students of an L2 are able to perceive intelligibility based on exposure to that L2. The same method can be applied to accents, as the longer an individual is in contact with a different regional accent, the speed and ease of understanding it is faster. This would explain why the Mid-Western North American and Southern North American accent speakers would have the highest engagement when it came to this accent. This explanation might also extend to the other two accents researched as well, Australian and London English, as this

would mean that a large chunk of the media coming out of the United States would also contain this same regional dialect.

On the whole, there was only a correlation between reaction time, accuracy, and nativeness of accent within one group of participants, the Mid-Western North American accent speaking participants, but delving further into other possible reasons for this occurrence could explain why this same accent was the most prevalent amongst the entire range of participant accent groups. This would be a great topic for further research.

## **6. Conclusion**

In conclusion, there appears to be little effect of nativeness on the speed and accuracy of perceiving emphasis in an accent of a person's mother tongue, at least based on the four accents studied. The findings of this experiment should be taken with some caution, as there were some pitfalls that could skew the results. Firstly, if this experiment is replicated in the future, it should be advised to employ more than the four accents presented in this study as this would expand the scope of the results by giving more data to analyze, as well as shedding light on how an even wider scope of accents interact with each other within the perception sphere. In a follow up experiment, a wider range of participants, stimuli and accents are necessary to draw anything more conclusive than what has already been stated in this paper. The effects of the results of this paper should inform people as to how they perceive an accent based on the media in which they consume it, and how that same media can control the world's view of a country's "native" accent. A cautionary tale that should be taken with a grain of salt.

## References

- Arnhold, Anja, Porretta, Vincent, Chen, Aoju (2020). (Mis)-understanding your native language: Regional accent impedes processing of information status. *Psychon Bull Rev* 27, 801–808. <https://doi.org/10.3758/s13423-020-01731-w>
- Arvaniti, Amalia, Garding, Gina (2007). Dialectal variation in the rising accents of American English. *Papers in laboratory phonology* 9, pp. 547– 576. Researchgate.net. Berlin:Mouton de Gruyter
- Baker, Caleb. (2020). Vowel length and word final voicing within English speech perception. Universiteit van Amsterdam. Unpublished
- Derwing, Tracey, Munro, Murray. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching*, 42(4), pp. 476-490. Cambridge University Press. doi:10.1017/S026144480800551X
- Deutskens, Elisabeth, de ruyter, ko, Wetzels, Martin, Oosterveld, Paul. (2004). Response Rate and Response Quality of Internet-Based Surveys: An Experimental Study. *Marketing Letters*. 15. 21-36. 10.1023/B:MARK.0000021968.86465.00.
- Floccia, Caroline, Goslin, Jeremy, Girard, Frédérique, Konopczynski, Gabrielle. (2006). Does a regional accent perturb speech processing?. *Journal of experimental psychology. Human perception and performance*. 32. 1276-93. 10.1037/0096-1523.32.5.1276.
- Herment, Sophie, Hirst, Daniel. (2002). Emphasis in English: a perceptual study based on modified synthetic speech. *Conference: 1st International Conference on Speech Prosody*, pp. 379-382
- Niedzielski, Nancy. (2002). Attitudes Toward Midwestern American English. *Handbook of Perceptual Dialectology*, pp. 323-329. John Benjamins Publishing. <https://doi.org/10.1075/z.hpd2>
- Reyes, Angela. (2017). Ontology of Fake: Discerning the Philippine Elite. Semiosis Research Center at Hankuk University of Foreign Studies. <https://doi.org/10.1086/690067>
- Thomas, Sarah, Scobbie, James. (2015). Mixed Accents: Scottish Children with English Parents. *Investigating English Pronunciation*, pp 83-104. Palgrave Macmillan, London. doi:10.1057/9781137509437\_4.
- Yan, Qin, Vaseghi, Saeed. (2003). Analysis, modelling and synthesis of formants of British, American and Australian accents. *IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. (ICASSP '03)*. doi: 10.1109/ICASSP.2003.1198880.



## Appendices

### Appendix 1: Surveys

#### Experimental

#### Participant instructions

##### For the Consent form

By checking the “Yes I agree” box, you agree to the consent form above.

##### For the experiment instructions

On the first slide, you will hear a phrase. After you have heard it, proceed to the next slide and of the four potential answers, say which word you believe was the most emphasized from the phrase before. Try to answer as fast as possible, as the second slide will be timed.

### Appendix 2: Stimuli

#### Movies used:

*The Good, the Bad, and the Ugly*. Leone, Sergio. Los Angeles, Calif.: Metro-Goldwyn-Mayer, 1967.

*Man of Steel*. Snyder, Zach. Burbank, CA: Warner Home Video, 2013.

*Crocodile Dundee*. Faiman, Peter. United States: Paramount Pictures, 1986.

*Crocodile Dundee II*. Bartelme, Jane Harper, Grant. United States: Paramount Pictures, 1988.

*Sherlock. Season one* [Television series]. London: BBC Worldwide, 2010

#### Target audio clips

##### Southern North American (*The Good, the Bad, and the Ugly*)

[https://www.youtube.com/watch?v=D7Ax5jr6mDM&ab\\_channel=Movieclips](https://www.youtube.com/watch?v=D7Ax5jr6mDM&ab_channel=Movieclips) (1.00 - 1.04)

[https://www.youtube.com/watch?v=rXs41JvueZY&ab\\_channel=Movieclips](https://www.youtube.com/watch?v=rXs41JvueZY&ab_channel=Movieclips) (0.28 - 0.32)

[https://www.youtube.com/watch?v=Qd9p6zTb2k0&ab\\_channel=MQ](https://www.youtube.com/watch?v=Qd9p6zTb2k0&ab_channel=MQ) (0.47 - 0.50)

##### Mid-western North American (*Man of Steel*)

[https://www.youtube.com/watch?v=tW8KyNCWGCY&ab\\_channel=CaptainDarrow](https://www.youtube.com/watch?v=tW8KyNCWGCY&ab_channel=CaptainDarrow) (1.10 - 1.12)

[https://www.youtube.com/watch?v=tEfoNXQDWBs&ab\\_channel=FlashbackFM](https://www.youtube.com/watch?v=tEfoNXQDWBs&ab_channel=FlashbackFM) (0.45 - 0.48)

[https://www.youtube.com/watch?v=tEfoNXQDWBs&ab\\_channel=FlashbackFM](https://www.youtube.com/watch?v=tEfoNXQDWBs&ab_channel=FlashbackFM) (1.28 - 1.31)

Australian (*Crocodile Dundee I & II*)

[https://www.youtube.com/watch?v=dSnosk4tWrg&ab\\_channel=TomasTree](https://www.youtube.com/watch?v=dSnosk4tWrg&ab_channel=TomasTree) (0.25 - 0.32)

[https://www.youtube.com/watch?v=2m2vmTtcCM8&ab\\_channel=Movieclips](https://www.youtube.com/watch?v=2m2vmTtcCM8&ab_channel=Movieclips) (0.40 - 0.43)

[https://www.youtube.com/watch?v=VmCRT88HTWE&ab\\_channel=Movieclips](https://www.youtube.com/watch?v=VmCRT88HTWE&ab_channel=Movieclips) (1.07 - 1.09)

English (*Sherlock*)

[https://www.youtube.com/watch?v=VaT7IYQgyqo&list=PL2qifpWIEmD605qsnCX4XwFt\\_7Z8P2hgr&ab\\_channel=Sherlock](https://www.youtube.com/watch?v=VaT7IYQgyqo&list=PL2qifpWIEmD605qsnCX4XwFt_7Z8P2hgr&ab_channel=Sherlock) (0.09 - 0.12)

[https://www.youtube.com/watch?v=exW90d6epQQ&list=PL2qifpWIEmD605qsnCX4XwFt\\_7Z8P2hgr&index=2&ab\\_channel=MikyAdler](https://www.youtube.com/watch?v=exW90d6epQQ&list=PL2qifpWIEmD605qsnCX4XwFt_7Z8P2hgr&index=2&ab_channel=MikyAdler) (0.11 - 0.13)

[https://www.youtube.com/watch?v=3DXh7nw2l0o&list=PL2qifpWIEmD605qsnCX4XwFt\\_7Z8P2hgr&index=3&ab\\_channel=RakitinaEkaterina](https://www.youtube.com/watch?v=3DXh7nw2l0o&list=PL2qifpWIEmD605qsnCX4XwFt_7Z8P2hgr&index=3&ab_channel=RakitinaEkaterina) (0.08 - 0.13)

**Filler audio clips**

Southern North American (*The Good, the Bad, and the Ugly*)

[https://www.youtube.com/watch?v=D7Ax5jr6mDM&ab\\_channel=Movieclips](https://www.youtube.com/watch?v=D7Ax5jr6mDM&ab_channel=Movieclips) (1.31 - 1.35)

[https://www.youtube.com/watch?v=rXs41JvueZY&ab\\_channel=Movieclips](https://www.youtube.com/watch?v=rXs41JvueZY&ab_channel=Movieclips) (0.15 - 0.19)

[https://www.youtube.com/watch?v=Qd9p6zTb2k0&ab\\_channel=MQ](https://www.youtube.com/watch?v=Qd9p6zTb2k0&ab_channel=MQ) (2.05 - 2.07)

Mid-western North American (*Man of Steel*)

[https://www.youtube.com/watch?v=tW8KyNCWGCY&ab\\_channel=CaptainDarrow](https://www.youtube.com/watch?v=tW8KyNCWGCY&ab_channel=CaptainDarrow) (0.56 - 0.58)

[https://www.youtube.com/watch?v=tEfoNXQDWBs&ab\\_channel=FlashbackFM](https://www.youtube.com/watch?v=tEfoNXQDWBs&ab_channel=FlashbackFM) (1.38 - 1.40)

[https://www.youtube.com/watch?v=tEfoNXQDWBs&ab\\_channel=FlashbackFM](https://www.youtube.com/watch?v=tEfoNXQDWBs&ab_channel=FlashbackFM) (1.10 - 1.13)

Australian (*Crocodile Dundee I & II*)

[https://www.youtube.com/watch?v=dSnosk4tWrg&ab\\_channel=TomasTree](https://www.youtube.com/watch?v=dSnosk4tWrg&ab_channel=TomasTree) (0.11 - 0.15)

[https://www.youtube.com/watch?v=2m2vmTtcCM8&ab\\_channel=Movieclips](https://www.youtube.com/watch?v=2m2vmTtcCM8&ab_channel=Movieclips) (0.29 - 0.32)

[https://www.youtube.com/watch?v=VmCRT88HTWE&ab\\_channel=Movieclips](https://www.youtube.com/watch?v=VmCRT88HTWE&ab_channel=Movieclips) (0.23 - 0.27)

English (*Sherlock*)

[https://www.youtube.com/watch?v=VaT7IYQgyqo&list=PL2qifpWIEmD605qsnCX4XwFt\\_7Z8P2hgr&ab\\_channel=Sherlock](https://www.youtube.com/watch?v=VaT7IYQgyqo&list=PL2qifpWIEmD605qsnCX4XwFt_7Z8P2hgr&ab_channel=Sherlock) (0.29 - 0.31)

[https://www.youtube.com/watch?v=exW90d6epQQ&list=PL2qifpWIEmD605qsnCX4XwFt\\_7Z8P2hgr&index=2&ab\\_channel=MikyAdler](https://www.youtube.com/watch?v=exW90d6epQQ&list=PL2qifpWIEmD605qsnCX4XwFt_7Z8P2hgr&index=2&ab_channel=MikyAdler) (0.49 - 0.51)

[https://www.youtube.com/watch?v=3DXh7nw2l0o&list=PL2qifpWIEmD605qsnCX4XwFt\\_7Z8P2hgr&index=3&ab\\_channel=RakitinaEkaterina](https://www.youtube.com/watch?v=3DXh7nw2l0o&list=PL2qifpWIEmD605qsnCX4XwFt_7Z8P2hgr&index=3&ab_channel=RakitinaEkaterina) (0.24 - 0.27)

**Appendix 3: R code****SNA LMER**

```
tableSNA$AccentN <- as.factor (tableSNA$AccentN)
contrast <- cbind (c(-1/2, +1/2))
colnames (contrast) <- c("native+non-native")
contrasts (tableSNA$AccentN) <- contrast

SNAlmer <-(lmer(Reactiontime~AccentN+(1|Participant), data=tableSNA))
```

**SNA GLMER**

```
tableSNA$Languagespeaker <- as.factor(tableSNA$Languagespeaker)
tableSNA$Accuracy <- as.factor(tableSNA$Accuracy)
tableSNA$AccentN <- as.factor (tableSNA$AccentN)
contrast <- cbind (c(-1/2, +1/2))
colnames (contrast) <- c("native+non-native")
contrasts (tableSNA$AccentN) <- contrast

SNAGlmer<-(glmer(Accuracy~AccentN+(1|Participant), family=binomial, data=tableSNA))
```

**SNA scatter plot**

```
library(ggplot2)
tableSNA$Languagespeaker <- as.factor(tableSNA$Languagespeaker)

ggplot(tableSNA, aes(x=AccentN, y=Reactiontime, colour = Languagelistener)) + geom_point()
+ geom_point(aes(x=1, y=14.702), colour="blue") + geom_point(aes(x=2, y=15.327),
colour="blue")+scale_y_continuous(breaks=c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20
))
```

**SNA stacked box plot**

```
library(ggplot2)
condition <- rep(c("Southern North American" , "Mid-Western North American" , "English" ,
"Australian") , 2)

ggplot(tableSNA, aes(fill=Accuracy, y=Languagelistener, x=Languagespeaker)) +
geom_bar(position="stack", stat="identity")
```

**MWNA LMER**

```
tableMWNA$AccentN <- as.factor(tableMWNA$AccentN)
contrast <- cbind(c(-1/2, +1/2))
colnames(contrast) <- c("native+non-native")
contrasts(tableMWNA$AccentN) <- contrast
```

```
MWNAImer <- (lmer(Reactiontime~AccentN+(1|Participant), data=tableMWNA))
```

**MWNA GLMER**

```
tableMWNA$AccentN <- as.factor(tableMWNA$AccentN)
contrast <- cbind(c(-1/2, +1/2))
colnames(contrast) <- c("native+non-native")
contrasts(tableMWNA$AccentN) <- contrast
```

```
MWNAglmer <- (glmer(Accuracy ~ AccentN+(1|Participant), family=binomial,
data=tableMWNA))
```

**MWNA scatter plot**

```
library(ggplot2)
tableMWNA$Languagespeaker <- as.factor(tableMWNA$Languagespeaker)

ggplot(tableMWNA, aes(x=AccentN, y=Reactiontime, colour = LanguageListener)) +
geom_point() + geom_point(aes(x=1, y=15.469), colour="blue") + geom_point(aes(x=2,
y=14.096), colour="blue") + scale_y_continuous(breaks = c(1,2,3,
4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20))
```

**MWNA stacked box plot**

```
library(ggplot2)
condition <- rep(c("Southern North American" , "Mid-Western North American" , "English" ,
"Australian") , 2)

ggplot(tableMWNA, aes(fill=Accuracy, y=LanguageListener, x=Languagespeaker)) +
geom_bar(position="stack", stat="identity")
```

**AUS LMER**

```
tableAUS$AccentN <- as.factor(tableAUS$AccentN)
contrast <- cbind(c(-1/2, +1/2))
colnames(contrast) <- c("native+non-native")
contrasts(tableAUS$AccentN) <- contrast
```

```
AUSlmer <- (lmer(Reactiontime~AccentN+(1|Participant), data=tableAUS))
```

**AUS GLMER**

```
tableAUS$Languagespeaker <- as.factor(tableAUS$Languagespeaker)
tableAUS$Accuracy <- as.factor(tableAUS$Accuracy)
tableAUS$AccentN <- as.factor(tableAUS$AccentN)
contrast <- cbind(c(0, 1))
colnames (contrast) <- c("native+non-native")
contrasts (tableAUS$AccentN) <- contrast
```

```
AUSglmer <- (glmer(Accuracy ~ AccentN+(1|Participant), family=binomial, data=tableAUS))
```

**AUS scatterplot**

```
library(ggplot2)
tableAUS$Languagespeaker <- as.factor(tableAUS$Languagespeaker)

ggplot(tableAUS, aes(x=AccentN, y=Reactiontime, colour = Languagelistener)) + geom_point()
+ geom_point(aes(x=1, y=15.293), colour="blue") + geom_point(aes(x=2, y=16.294),
colour="blue") + scale_y_continuous(breaks = c(1,2,3,
4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20))
```

**AUS stacked box plot**

```
library(ggplot2)
condition <- rep(c("Southern North American" , "Mid-Western North American" , "English" ,
"Australian") , 2)

ggplot(tableAUS, aes(fill=Accuracy, y=Languagelistener, x=Languagespeaker)) +
geom_bar(position="stack", stat="identity")
```

**E LMER**

```
tableE$AccentN <- as.factor(tableE$AccentN)
contrast <- cbind(c(-1/2, +1/2))
colnames (contrast) <- c("native+non-native")
contrasts (tableE$AccentN) <- contrast
```

```
Elmer <- (lmer(Reactiontime~AccentN+(1|Participant), data=tableE))
```

**E GLMER**

```
tableE$Accuracy <- as.factor(tableE$Accuracy)
tableE$AccentN <- as.factor(tableE$AccentN)
contrast <- cbind(c(-1/2, +1/2))
colnames (contrast) <- c("native+non-native")
contrasts (tableE$AccentN) <- contrast
```

```
Eglmer <- (glmer(Accuracy ~ AccentN+(1|Participant), family=binomial, data=tableE))
```

### **E scatter plot**

```
library(ggplot2)
tableE$Languagespeaker <- as.factor(tableE$Languagespeaker)

ggplot(tableE, aes(x=AccentN, y=Reactiontime, colour = LanguageListener)) + geom_point() +
geom_point(aes(x=1, y=11.875), colour="blue") + geom_point(aes(x=2, y=11.905),
colour="blue") + scale_y_continuous(breaks = c(1,2,3,
4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20))
```

### **E stacked box plot**

```
library(ggplot2)
condition <- rep(c("Southern North American" , "Mid-Western North American" , "English" ,
"Australian") , 2)

ggplot(tableE, aes(fill=Accuracy, y=LanguageListener, x=Languagespeaker)) +
geom_bar(position="stack", stat="identity")
```