Mānoa Horizons

Volume 3 | Issue 1 Article 8

11-15-2018

A Speech Rhythm Comparison of English Speakers from Hawai'i and California

Lana Lobato Aljader
University of Hawai'i at Mānoa

A Speech Rhythm Comparison of English Speakers from Hawai'i and California

Lana Lobato Aljader

Honors Thesis (Interdisciplinary Studies: Communication Science and Disorders)
Mentors: Dr. Victoria B. Anderson and Dr. Katie Drager

There has been a long tradition, since at least 1945, of research into the rhythm of speech. Considered a universal feature of language, speech rhythm is often broken down into two main categories, stress-timed and syllable-timed. Languages are assumed to fit into, or fall along a continuum between, these two categories. This study compares the recorded speech of two politicians speaking Californian English, a so-called stress-timed language, and two politicians speaking Hawai'i English, which has yet to be categorized. Influencing languages, regional dialect, and social impacts are discussed. The software DARLA and the program Praat were used to assist in the manual insertion of vowel boundaries. To compare the ratio of differences in duration between successive vowels for each speaker the normalized Pairwise Variability Indices (nPVI) was calculated. The hypothesis that Hawai'i English is more syllable-timed than other American varieties, as has been impressionistically observed in the literature, was not supported. Limitations of the traditional conception of rhythm and of the current study, as well as the need for further work, are discussed.

1. Introduction

The following introductory sections cover research into speech rhythm and the provenance of Hawai'i English. The foundational idea that language can be categorized as syllable-timed or stress-timed is discussed, as well as critiques of this theory. Section 1 also discusses regional dialects and other related studies relevant to this research. Section 2 summarizes the methods used in this study. Section 3 states the results for each sample and section 4 discusses those findings. Section 5 concludes the paper. The first thing that will be discussed in this introduction is speech rhythm.

1.1 Speech Rhythm

When we hear the word *rhythm*, speech is not the first thing one typically thinks of. Most think of rhythm in relation to music; however, rhythm can also be found in our speech. Rhythm is a feature of prosody in speech that differs between languages

I am a recent graduate of the University of Hawai'i at Mānoa with a degree in Interdisciplinary Studies: Communication Science and Disorders. This research is a revised version of my Honors undergraduate thesis. This project was a collaboration of the field of CSD and Linguistics through the aid of my mentor Dr. Victoria B. Anderson and committee member Dr. Katie Drager. I hope to continue my education by completing a master's degree and becoming a licensed Speech Language Pathologist.

Mānoa Horizons, Vol. 3, 2018, pp. 36–43

and regions. Rhythm is a widely accepted feature of language. However, distinctions in types of rhythm among languages have been hypothesized, challenged, and developed over time.

Language is a complex feature that plays an important role in comprehension. Rhythm plays a crucial role in a listener's ability to process and comprehend human speech (Cutler, 1991). It has been widely agreed upon that a listener actively processes the rhythm of someone's speech and subconsciously learns to anticipate the patterns as they occur. This is important in the comprehension of language because if the speech suddenly disrupts the anticipated rhythmic pattern, the listener will have a momentary lapse of recognition and understanding (Darwin, 1975, Meltzer et al., 1976, Martin, 1976, Buxton, 1983).

The most widely accepted categorization of linguistic rhythms was developed by Kenneth Lee Pike in 1945. Pike proposed a binary system of categorization. Languages were either stress-timed or syllable-timed. Stress-timed languages have longer durations for stressed syllables than unstressed syllables. Syllable-timed languages have syllable durations similar to each other throughout speech, also known as isochrony (Pike, 1945, p. 72). Pike found English and Spanish to be the prototypical languages for stress-timing and syllable-timing, respectively. Further research showed that syllable-timed prototypes were also French (Abercrombie, 1967, Ladefoged, 1975, Delattre, 1951), and Japanese (Ladefoged, 1975).

1.2 Critiques of Stress-timed vs. Syllable-timed Categorizations

Although the categorization of languages as either syllable-timed or stress-timed was initially met with acceptance, there has been much research on the validity of this categorical system. A main point that many critics focus on are the two definitions of the categories laid out by Abercrombie:

1. 'there is considerable variation in syllable length in a language spoken with stress-timed rhythm whereas in a language spoken with a syllable-timed rhythm the syllables tend to be equal in length'.
2. 'in syllable-timed languages, stress pulses are unevenly spaced'. (Abercrombie, 1967).

Roach (1982) hypothesized that the differences found between languages were not actually due to syllable duration, but to other features. He came to this conclusion in his analysis of French, Telugu, Yoruba, English, Russian, and Arabic when he found no clear distinction in syllable timing. These findings conflicted with previous research done on French, Telugu, and Yoruba which had been accepted as prototypical syllable-timed languages (Abercrombie, 1967). Flaws in the categorization began to emerge.

The differences in results can be attributed to the varying types of methods used across linguistic practices. Pamies Bertrán highlighted these differences in 1999, "Navarro Tomás uses a poem by Rubén Dario, Gili Gaya, the reading of an essay, Delattre and Olsen, spontaneous speech, while the bulk of work by Toledo (1988) includes spontaneous speech, essay, literary prose or poems read by the authors (Ernesto Sábato, García Márquez, Nicolás Guillén)" Bertrán's work displays the variety of potential influences on syllable duration that make it a challenge to have one accepted form of methodology across the field for more accurate results.

Due to research which revealed clear flaws in the binary approach of categorization, there was a move towards understanding rhythm as a continuum, with stress-timed and syllable-timed at each end, where languages land somewhere along the continuum (Ramus et al, 1999, Dauer, 1987). The current study hypothesized that speakers from the State of California would be closer to the stress-timed end of the continuum, whereas speakers from the State of Hawai'i would be closer to the syllable-timed end of the spectrum.

1.3 Regional Dialects

Before looking further into the current study, it is important to understand where the languages we are comparing came from. We will be looking at samples from the State of Hawai'i and the State of California. The region of California was used to represent Standard American English (SAE). Hinton et al. (1987) found that speakers in California did not have a regionally specific dialect, and there were no regionally specific vowel qualities to that area that differentiated it from with the rest of the country. So Californian speakers will be used as an example of SAE.

While there is literature on Californian English (Hinton et al., 1987), there is far less published material on Hawai'i English. For the purposes of this study, Hawai'i English is being treated as a regional dialect of the State of Hawai'i within the English speaking United States of America (Drager, 2012; Drager et al., 2013). However, for background on HE, a discussion of the origin and structures of Hawaiian Creole, a variety that strongly influences HE, is necessary.

1.3.1 Hawaiian Creole

Hawaiian Creole (HC) originated as a form of communication in the Hawaiian Islands in the mid to late 1800's, as a result of the large immigration of plantation workers. As the sugar cane plantation industry grew in Hawai'i so did its need for laborers. Most of the immigrants did not speak English and originated from different countries around the world. In order to communicate with each other, a pidgin language developed in the region, which later devolved into a creole language. What is now known as Pidgin, Hawaiian Creole English, or Hawaiian Creole, is based off of English, Chinese, Portuguese, Japanese, Hawaiian, and Filipino languages (Marlaw & Giles, 2008). For years, Hawaiian Creole was viewed as unacceptable to use in

formal speech. As recently as 2008, Marlaw and Giles found that "although they [residents of Hawai'i] speak Pidgin on a routine basis in social, familial, and employment contexts (i.e., with other employees rather than tourists or bosses), they speak Standard English in order to conform to professional norms of appropriateness (often learned by parents and family) during employment interviews, interactions with supervisors or in educational contexts" (Marlaw & Giles, 2008). Since Marlaw and Giles' study was published, HC became recognized as an official language of the State of Hawai'i in 2015. A language that was once criticized and unaccepted formally is now making its way beyond people's homes and social interactions into more aspects of their daily lives.

1.3.2 Linguistic characteristics of Hawaiian Creole and Hawai'i English

The terms "Hawai'i English" (HE) and "Hawaiian Creole" (HC) were established by Sato in 1993. While English has been found to be the main lexifier of HC (Drager et al., 2012, 2013), HC and English have many linguistic differences and similarities.

First, we see different patterns of stress. Most two syllable words that have origins from Standard American English (SAE) have stress in the same place for both languages. However, there are exceptions. These exceptions are found in words such as "Beis**bawl** (**base**ball), chap**stik** (**chop**stick), and dead**lain** (**dead**line)" (Sakoda, 2003), where the bold shows the stress of the word. It has also been found that HC has a different primary stress when a word has two or more syllables. These words are found to usually end in –ary, -ony, or –ory. HC has a tendency to have more stress on syllables overall than SAE does, and more equally prominent syllables.

Even though HC has been highly influenced by SAE's lexicon, it has attributes derived from the other languages of influence. A main difference between SAE and HC is the placement of adjectives and stative verbs can be found before the noun. This placement is shown in phrasing such as

a.) 'Ono ka i'a. (Hawaiian)
delicious DET fish
'The fish is delicious.'

b.) Cute da baby. (Pidgin) (Drager, 2012).

Looking at HE and HC together, we see that these speech varieties have their own characteristics that make them distinct from dialects of the continental United States. Drager and colleagues found that there were differences between HC and SAE speakers with respect to vowel quality (Drager, 2012, Drager et al., 2013). The most relevant finding for the current study is that HC is closer on the continuum to syllable-timed languages than SAE (Sakoda & Siegel, 2003).

1.3.3 Social influence on HC

There is limited research on the linguistic characteristics of Hawaiian Creole, due to social mentalities regarding the language. "Historically, Pidgin English was basically only used by the working class—whom the upper-class Caucasians felt spoke an inferior form of English spoken by the ignorant and uneducated. This stereotype continues to be present, and the current friction between Pidgin English proponents and opponents has its roots in Hawai'i's social inequalities" (Kawamoto, 1993:201, cited in Higgins et al., 2010). The use of HC particularly in professional or academic settings was continuously challenged. In 1987, legislation was drafted that would ban children from speaking HC in the classrooms. But due to public outrage the bill was stopped. They were, however, able to pass a bill encouraging teachers to only allow English in their classroom (Drager, 2012; Drager et al., 2013).

Similar discrimination was met by professionals in the community, specifically those in the public eye. In the same year of 1987, a lawsuit involving the National Weather Service was established under the premise of linguistic profiling (Drager, 2012; Drager et al., 2013). The case of Kahakua et al. v Hallgren involved two meteorologists from Hawai'i suing the National Weather Services because they were not hired due to the fact they spoke HC (Higgins et al., 2010). The company instead hired two people with less experience because they sounded better to the National Weather Service. During the hearing, the judge ruled against the local plaintiffs and is quoted as saying they should "Put more effort into improving their speech" (Sato, 1991:655, cited in Higgins et al., 2010).

Although rhythmic differences are able to be established as categories and numerical distinctions the effects or differences may also have cultural implications that reflect identity. Differences in regional dialects and their distinctions including speech rhythm, may reflect a sense of identity and be attributed to a sense of belonging and ancestry. In the case of HE's origins, for those of native blood the Hawaiian language is sacred and its progression of melding with English and other languages to form HC are an important part of history and heritage. As we look closer at HE by trying to recognize these differences and the effect that Hawaiian and HC have on English it could give a further sense of pride and belonging for a group of people that have been marginalized and whose language has been rejected by their country and fellow people throughout history, as disused above in the case of Kahakua et al. v Hallgren. However, by showing the rhythmic influences of minority language in politicians this may support growth of the status of the population. The importance of regional dialects and distinctions within this study is further discussed with the state of California in section 4.1.

1.4 Studies of Rhythm in English Dialects

Previous research has looked at timing differences among regional dialects within English. Three studies were done looking at Singapore English Pronunciation (SEP). It was believed that SEP, though a dialect of English, was syllable-timed rather than stress-timed (Deterding, 1994). When measuring the durations of stressed and unstressed syllables, Yeow (1987) did not find supporting evidence. However, Low (1994) did find supporting evidence when measuring the variability of vowel durations. Low's results were supported by Deterding (1994) in comparison to Southern Standard British English. These results are relevant to the current study as they show it is possible to have a categorically English dialect that developed to be closer to a syllable-timed language on the spectrum, due to other influences. Deterding believes these different features are due to "final consonant cluster simplification, the relative absence of vowel reduction in unstressed syllables, and the lack of smooth liaison between words" (Deterding, 1994).

2. Methods

For this study four subjects were chosen for analysis: two from Hawai'i and two from California. To control for variability that may be caused by gender differences, only male subjects were selected to be analyzed. From each of these four male participants a formal speech was chosen. The four participants are all local politicians with roots in the area they are representing represented by birthplace, education, and residence. The two representatives selected from Hawai'i were David Ige and Shan Tsutsui. The two representatives selected from California were Jerry Brown and Eric Garcetti.

2.1 Speech Materials

Various formal public speeches by the four politicians were reviewed over the format of YouTube. Speeches needed to fit the requirements of a ten-minute time length and be presented in front of an audience in a formal setting to control for speech style. For the Hawai'i English speakers, Governor David Y. Ige's 2014 inauguration speech was chosen along with Lieutenant Governor Shan Tsutsui's address regarding the proclamation for fire prevention week. For the California representatives, Governor Jerry Brown is addressing the Oroville Spillway response and recovery from the State Operations center and the final sample is from Los Angeles Mayor, Eric Garcetti's 2013 inauguration speech. These samples were converted into MP3 files. Two minutes after the beginning of each speech, a five-minute section was selected for measurement and analysis.

2.2 Speech Segmentation

Once the speech samples were selected, the goal was to determine the relative isochrony between successive vowels. The MP3 files were run through the Dartmouth Linguistic Automation (DARLA) segmenter's Completely Automated Vowel Extraction (Reddy & Stanford, 2015). The result was two types of automated transcription: words aligned with the audio, and segments aligned with audio. These transcriptions were the starting point for the hand-adjusted segmentation of vowels for the selected samples.

Once the transcription files for Governor Ige, Lieutenant Governor Tsutsui, Governor Brown, and Mayor Garcetti were received they were matched with their respective MP3 files and opened with the software Praat (Boersma & Weenink, 2017). This program allows MP3 files to be shown in spectrogram and waveform view along with accompanying TextGrids. The program has various functionalities in relation to phonetic analysis; however, for this research Praat was mainly used to hand mark vowel lengths. The MP3 file and TextGrid file supplied by DARLA were simultaneously opened and the function "view & edit" was selected. The aligned waveform, spectrogram, and TextGrids labeling the vowels' boundaries and the transcript of speech were generated. The starting point was made around 120 seconds or later. This selection was to ensure that the speaker had become comfortable and reached a consistent speaking rate. The end point was made 300 seconds after the starting mark for each sample. Thus, with four samples, a total of 1,200 seconds of audio were analyzed. Using the visual aid of the spectrogram, waveform, and formants, each of the automated vowel extraction labels was checked by hand to ensure accuracy of start and end points. An example is shown in figure 1 below.

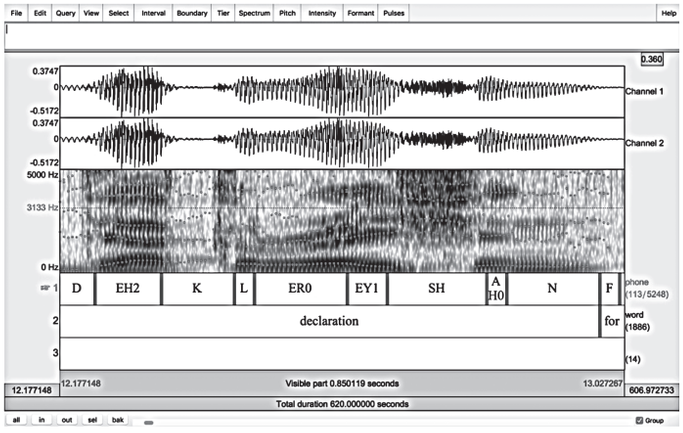

Figure 1 Praat display of the spectrogram, waveform, formants, and automated vowel extraction labels.

2.2.1 Automatic duration extraction

A Praat script was then run (Workshops & Tutorials, 2018) to extract the durations of all labeled segments in a TextGrid object. The script was previously downloaded and saved as a file on the computer in use. Back at the Praat Objects window the TextGrid file was selected. Under "Praat" the option of "Open Praat script" was then chosen. The saved script file was selected, and the script immediately opened in a separate window. The button "Run" was chosen and the specific Tier was input and the button "Okay" selected. A file was saved onto the computer with each segment and its respective duration. This process was repeated for each audio sample TextGrids.

2.3 Data Analysis

Following the method of Ling et al. (2000), consecutive vowel durations were compared in the four speech samples, to determine relative isochrony or variability between successive vowels. The measurements were also normalized for differences in speaking rate. The metric that was used is called the normalized Pairwise Variability Index (nPVI). This method was first introduced by Ling and associates in 2000, and the method for calculating it is shown in Figure 2. The nPVI calculates the absolute value of the difference between successive pairs of vowels ([dk-dk+1] in the formula below), and then normalizes for speech rate by dividing the difference by the mean duration of the pair ([dk+dk+1]/2 in the formula below), and then sums up all the differences, and divides by the number of differences (m-1) in order to create a numerical representation. This total is then multiplied by 100 to give a whole number representation.

For each speech sample, durations of the labeled TextGrids were entered into an Excel sheet, and any non-vowel segments were deleted. The order of the vowels in the audio sample were maintained. The raw vowel durations were then inputted into an nPVI Excel spreadsheet provided by Ladefoged and Johnson (2005), resulting in a different nPVI for each speaker. For example, in Figure 3, Ladefoged and Johnson compare nPVI values for nine different languages. Note that the nPVI for English is about 55. From this score, we can hypothesize that for the current study, the speakers from the State of California will have a score around 55 and the speakers from the State of Hawai'i will have nPVIs that are lower.

$$\mathbf{nPVI} = 100 \times \left(\sum_{k=1}^{m-1} |(d_k - d_{k+1})/((d_k + d_{k+1})/2)| \right) \Big/ (m-1),$$

Figure 2 Calculation of the normalized Pairwise Variability Index.

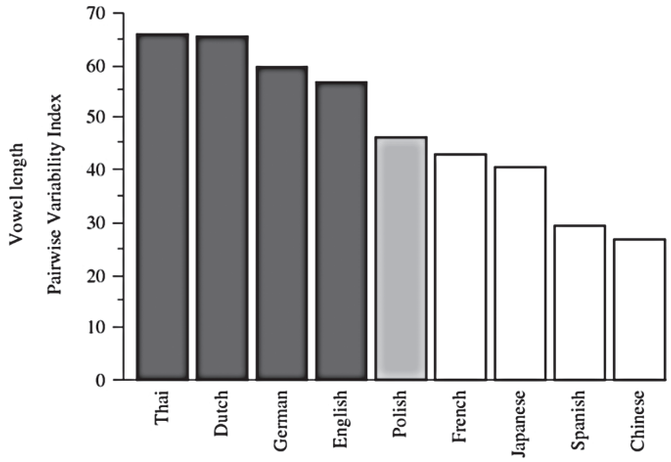

Figure 3 The nPVI scores of 9 different languages (Ladefoged and Johnson, 2005).

3. Results

The final four nPVI scores for each subject's speech are found in Table 1. The Hawai'i politician David Ige had a final nPVI score of 49.5. Shan Tsutsui had a final score of 68.2. Eric Garcetti received a score of 61. Jerry Brown had a score of 74.2. These results are also shown in Figure 4 with the Hawai'i politicians' results shown in green and the California politicians' results shown in orange.

The original hypothesis of this study was that the Hawai'i English speakers would have lower nPVI scores than the California English speakers. However, this study showed mixed results. From lowest to highest the politicians' nPVIs were ranked as Ige (Hawai'i), Garcetti (California), Tsutsui (Hawai'i), and Brown (California). Ige followed the predicted results of being lower than the English average of 55 nPVI (Ladefoged & Johnson, 2005), however Tsutsui scored above that margin by 13.2. Moreover, while Garcetti and Brown both scored above 55, perhaps the most important result is that Garcetti's Californian English had a lower nPVI than Tsutsui's Hawai'i English. The original hypothesis was not supported by these results.

Table 1 nPVI scores of the 4 politicians.

POLITICIAN	NPVI
IGE	49.5
TSUTSUI	68.2
GARCETTI	61
BROWN	74.2

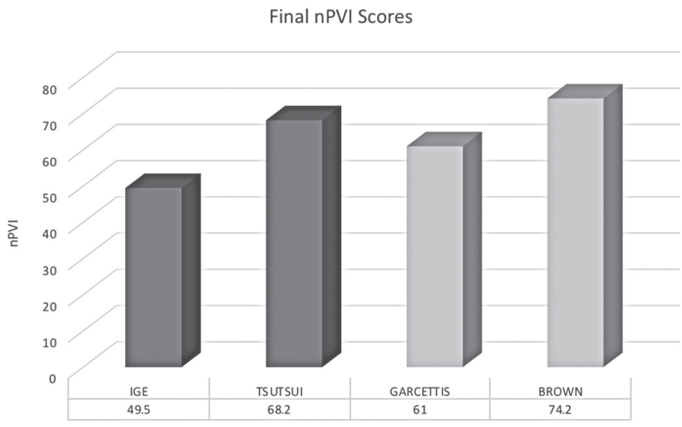

Figure 4 nPVI scores of the 4 politicians in graphical form.

4. Discussion

The remainder of this section will address three types of reasons that might explain why the results did not bear out the original hypothesis. These are idiosyncratic influences from other languages, the nPVI measure itself, and the idea from Nolan and Jeon (2014), that speech rhythm may be something that is not actually measurable by metrics like nPVI.

4.1 Influences of other speech varieties

This study looked at a small group of four samples. For a better understanding of the study's question more participants would need to be analyzed to create a larger sample size of data. The results could have also been affected by differences between the speakers that were not accounted for. A deeper look into their backgrounds and the backgrounds of their parents could give more information to their results. For the instance of Garcetti, he was born and raised in Southern California. There is a large presence of the Spanish language in the region which is also a so-called prototypical syllable timed language (Pike, 1945:72). In 1990 there were 8.6 million non-English language speakers or 32% of the state, and Spanish was the predominant alternative language in California (Marcias, 2001). A future question to look into would be whether the presence of Spanish in his environment had an effect on Garcetti's score. Garcetti has a mixed ancestry that includes a grandfather from Mexico. Although Garcetti is said to be fluent in Spanish, Fausset (2014) refers to it as "Funky American Business Spanish." The mayor learned his Spanish mainly through academics and business (Fausset, 2014; Jamison, 2016).

Research done by Bucholtz, et al. (2007) prompted 70 students attending the University of California, Santa Barbara in a sociolinguistics course with a blank map of California and the directions "Please draw a boundary around each part of California where you believe people speak differently, and label the area" (Bucholtz, et al., 2007). The results showed that the students marked the Southern California region, particularly San Diego, as "Mexican," representing the heavy influence of people from Mexico and the Spanish language. Northern California was commonly marked as "normal" or "standard" English while Southern California was not. Therefore, regional discrepancies between Garcetti and Brown could have been a factor in their nPVI scores. Further research would be needed to support those claims.

A future study that could be used to further understand the differences between the speech produced in the two states would be to conduct a pre-survey to help determine the samples selected. In the state of Hawai'i it is not uncommon to hear the phrase "you sound local" or the opposite, "you don't sound local." If an initial survey was given to select samples that the local population perceived as representing what is viewed as local Hawai'i English speech, clearer differences might appear. This study was based off the determining factors of the location of birth place, education, and professional careers. However, that may not have been enough to determine proper representations of the targeted dialects. Languages spoken by family members and more information about the parents and upbringing of the sample speakers could be necessary for selection.

4.2 Problems with nPVI

A potential problem with the methodology of this study could be the use of the nPVI score itself. First, the choice of boundary locations could strongly affect nPVI results. Human error and irregularity could cause large differences in results. "More fundamentally, it remains to be demonstrated whether such metrics really quantify speech rhythm, a controversial and elusive concept" (Kohler, 2009).

These problems could be improved by widening the understanding of rhythm. Speech rhythm should account for more than just duration measurements. Pitch may have as much effect on rhythm as timing. In a study done by Nolan & Jeon in 2014 comparing Swiss-German and Swiss-French the F0 was manipulated and participants were asked to choose which was rhythmically more natural. Syllables were changed in duration and either increased or decreased by 35%. Results found that the Swiss-German speakers were not accepting of durational manipulation of the crucial syllables as rhythmically sounding "natural." However, they did find variation in pitch range as appearing to be appropriate. In contrast, the Swiss-French speakers chose shorter duration and smaller pitch range as appropriate speech but were not as accepting of longer duration and increased pitch ranges. "Nonetheless, the experiment shows that where the task is to judge rhythm, subjects are sensitive to pitch as well as timing, and, less predictably, that the integration of these dimensions may well be language-specific." (Nolan & Jeon, 2014).

A study done by Arvaniti in 2009 found that "the elicitation method can significantly affect metric scores in such a way

that rhythmic classifications can change depending on the type of data on which scores are calculated." The material selected can cause a sample expecting to be syllable-timed to produce a score more closely representing a stress-timed language. Arvaniti created phrases that were either constructed as simple consonant-vowel alterations (producing an optimally 'syllable-timed' version of that language) and a phrase constructed with the most segmental variation possible for the respective language (producing an optimally 'stress-timed' language). They found that by intentionally altering the samples in that way there were differences in their scores and language classifications. The variability in samples used within a language can have an effect on how it is classified, producing further complications in the reliability of languages being able to categorize rhythmically (Arvaniti, 2009).

One of the challenge in measuring speech and rhythm is a large between-speaker variability. The speakers being tested can have varying styles that can cause large differences, and the speech "task" (such as reading a story, reading sentences, or spontaneous speech) can affect measures. Focusing on a concept of rhythm that does not rely on grouping and patterns, but only on strict measures of timing, can cause error. Another aspect of concern is cross-linguistically valid results because syllable-timed and stress-timed dichotomies were created by English speakers with the English language's view and perception of rhythm. These potential problems in results could be helped by a change in the attention to the range of what is studied in speech rhythm. Arvaniti (2009) discusses that a potential solution to further study focuses on the idea that stress should not be viewed as the basis of rhythm in all languages but perhaps to categorize languages from least to most stress based instead (Nolan & Jeon, 2014).

Other features to look closer at are the prominence gradient, a range of segmental and suprasegmental factors. Items that may influence prominence include "intrinsic segmental duration, syllable structure, lengthening or shortening at the edge of word or prosodic unit, durational adjustment related to the distribution or relation of stress or accent, and speech rate, not solely based on the language's 'rhythm unit' " (Nolan & Jeon, 2014). Accepting these variabilities and using a combination of acoustic cues rather than just duration may create the proper circumstances to produce a more accurate representation of rhythm (Kohler, 2009; Nolan, 2014).

4.3 Measurability of Speech Rhythm

Nolan and Jeon (2014) separate rhythm into two categories: Coordinative or Contrastive. Coordinative rhythm means regularity in time. Nolan and Jeon (2014) compare this distinction to the sounds of a saw moving back and forth, or a train on tracks. In speech that difference would be reflected in isochrony of units; repeated elements or groups would occur with the same time span. However, contrastive rhythm is defined by a shift between strong and weak elements. These elements do not need to be tied to an external clock such as a metronome like Coordinative rhythm does. Nolan and Jeon suggest that language itself could be arhythmic. The fact that there are conflicting definitions of rhythm brings into question the legitimacy of measuring rhythm in speech as a whole, and how empirical the process can be.

Kohler also brings up the question of whether rhythmic differences are due to personal perception and production. Rhythm may be a matter of performance quality as a tool used to guide a listener through different meanings. Kohler (2009) uses figures such as Barack Obama or Martin Luther King as examples; compared with others in society there appears to be strong rhythmic difference in their styles of speech. Further research is needed to understand if our results have been a reflection of the rhythmic differences among individuals instead of our previously accepted categorical distinctions (Nolan & Jeon, 2014).

5. Conclusion

This study used a traditional method of attempting to measure rhythm based on the foundational works of Pike (1945) and those building on the topic thereafter. The fact that language falls on a continuum between stress-timed and syllable-timed languages was assumed. The question was posed whether political figures in Hawai'i had a difference in speech rhythm in comparison to the political figures in California. This study put that to the test by comparing successive vowels based on their normalized Pairwise Variability Index scores. It was hypothesized that Hawai'i English speech would be closer to syllable-timed than Californian English speech. Results did not clearly divide these speakers into groups based on their nPVI scores alone.

The results of this study may have been the result of flaws in the methodology related to vowel segmentation, influences of other speech varieties, problems with the normalized Pairwise Variability Index itself, and the validity of measuring speech rhythm in production without considering perception. The definition of rhythm may need to be revisited and the legitimacy of how it is measured may need to be called into question, expanding what rhythm means to include how it is perceived, not just produced, and including different cues like pitch, and amplitude. Further work on the subject is necessary to have a better understanding of what rhythm is, its relation and applicability to speech, and how it can be properly and accurately measured and represented.

References

Abercrombie, D. 1967. *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

Arvaniti, A. (2009). Rhythm, timing and the timing of rhythm. *Phonetica*, *66*(1–2), 46–63.

Boersma, Paul & Weenink, David (2017). Praat: doing phonetics by computer [Computer program]. Version 6.0.20, retrieved 6 September 2016 from http://www.praat.org/

Bucholtz, M., Bermudez, N., Fung, V., Edwards, L., & Vargas, R. (2007). Hella Nor Cal or totally So Cal? the perceptual dialectology of California. *Journal of English Linguistics*, *35*(4), 325–352.

Buxton, H. (1983). Temporal predictability in the perception of English speech. In Cutler, and Ladd, D.R. (eds.), *Prosody: Models and Measurements*, Springer, Heidelberg.

Cutler, A. (1991). Linguistic rhythm and speech segmentation. In *Music, language, speech and brain* (pp. 157–166). Macmillan Education UK.

Darwin, C.J. (1975). On the dynamic use of prosody in speech perception. In Cohen, A. and Nootcboom, S.G. (eds.), *Structure and Process in Speech Perception*, Springer, Berlin.

Dauer, R.M. (1983). Stress-timing and syllable-timing reanalyzed. *J. Phon.,* 11, 51–62.

———. (1987). Phonetic and phonological components of language rhythm. *Proc.11th Int. Cong. Phon. Sci.t* Tallinn; Vol. 5

David Y. Ige. (n.d.). Retrieved from http://governor.hawaii.gov/governors-bio

Delattre, P. (1951). The physiological interpretation of sound spectrograms. *Publications of the Modern Language Association of America*, 864–875.

Deterding, D. (1994, December). The rhythm of Singapore English. In *Paper presented at Fifth Australian International Conference on Speech Science and Technology* (Vol. 6, p.8).

Drager, K. (2012). Pidgin and Hawai 'i English: an overview. *International Journal of Language, Translation and Intercultural Communication*, *1*(1), 61–73

Drager, K., Kirtley, M. J., Grama, J., & Simpson, S. (2013). Language Variation and Change in Hawai'i English: KIT, DRESS, and TRAP. *University of Pennsylvania Working Papers in Linguistics*, *19*(2), 6.

Fausset, R. (2014, March 04). In Mexico, L.A. Mayor Eric Garcetti's Spanish goes far. *Los Angeles Times*. Retrieved from http://articles.latimes.com/2014/mar/04/world/la-fg-mexico-mayor-20140305

Higgins, C., Léglise, I., & Bartens, A. (2010). Raining critical language awareness in Hawai'i at Da Pidgin Coup. In B. Miggle (Ed.), *Creoles in Education: An appraisal of current programs and projects* (Vol. 36, Creole Language Library, pp. 31–54). Amsterdam/Philadelphia: John Benjamins Publishing.

Hinton, L., Moonwomon, B., Bremner, S., Luthin, H., Van Clay, M., Lerner, J., & Corcoran, H. (1987, September). It's not just the Valley Girls: A study of California English. *In Annual Meeting of the Berkeley Linguistics Society* (Vol. 13, pp. 117–128).

Jamison, P. (2016, July 28). L.A. mayor Eric Garcetti introduces himself to nation at DNC: 'I'm just your average Mexican American Jewish Italian'. *Los Angeles Times*. Retrieved from http://www.latimes.com/politics/la-na-pol-garcetti-convention-20160728-snap-story.html

Kohler, K. J. (2009). Whither speech rhythm research?. *Phonetica*, *66*(1–2), 5.

Ladefoged, P., & Johnson, K. (2005). *A Course in phonetics* (Fifth ed.). Stamford, CT: Cengage Learning.

Low, E. L. (1994) *Intonation Patterns in Singapore English*. M.Phil. Dissertation, Department of Linguistics, Cambridge University.

Ling, L. E., Grabe, E., & Nolan, F. (2000). Quantitative characterizations of speech rhythm: Syllable-timing in Singapore English. *Language and speech*, *43*(4), 377–401.

Marlaw, M. L., & Giles, H. (2008). Who You Tink You, Talkin Propah? Hawaiian Pidgin Demarginalised. *Journal Of Multicultural Discourses*, *3*(1), 53–68. doi:10.1080/17447140802153535

Martin, J.G. (1979). Rhythmic and segmental perception are not independent. *J. Acoust, Soc, Amer,,* 65, 1286–1297.

Meltzer, R.H., Martin, J.G., Mills, C.B., Imhoff, D.L. and Zohar, D. (1976). Reaction time to temporally displaced phoneme targets in continuous speech. *J. Exp. Psy,: Hum, Percf Perf.,* 2, 277–290.

Nolan F, & Jeon H-S. (2014). *Speech rhythm: a metaphor?* Phil. Trans. R. Soc. B 369 : 20130396. http://dx.doi.org/10.1098/rstb.2013.0396

Pamies Bertrán, A. (1999). Prosodic typology: on the dichotomy between stress-timed and syllable-timed languages. *Language design: journal of theoretical and experimental linguistics*, *2*, 103–130

Pike, K.L. 1972 [1945]. "The intonation of American English". In Bolinger, D. (ed.): *Intonation*. Harmondsworth: Penguin: 53–83.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, *73*(3), 265–292.

Reddy, Sravana and James Stanford. 2015. Toward completely automated vowel extraction: Introducing DARLA. *Linguistics Vanguard*.

Roach, P. (1982). On the distinction between 'stress-timed'and 'syllable-timed'languages. *Linguistic controversies*, *73*, 79.

Sakoda, K., & Siegel, J. (2003). *Pidgin Grammar: An Introduction to the Creole English of Hawai'i.* Honolulu, HI: Bess Press.

Sato, C. J. (1993). Language change in a creole continuum: Decreolization. *Progression and regression in language: Sociocultural, neuropsychological, and linguistic perspectives*, 122–143.

Shan S. Tsutsui. (n.d.). Retrieved April 01, 2017, from http://ltgov.hawaii.gov/about/lieutenant-governors-biography/

Yeow, K. L. (1987) *Stress, Rhythm and Intonation in Educated Singaporean English*. M.A. Dissertation, National University of Singapore.

Workshops & Tutorials. *LAELabs*. (n.d.). Retrieved 2018, from https://sites.google.com/a/hawaii.edu/laelabsedit/workshops