A cross-generational investigation of voice quality in women

Sarah D. F. Greer, Stephen J. Winters
University of Calgary

Abstract

This research investigates the use of creaky voice by university-aged women and their mothers in order to answer three main questions: i. is there a specific phonetic environment where this voice quality is more likely to occur, ii. do young women use this voice quality more frequently than older women?, and iii. is creaky voice a register marker? Five mother-daughter pairs were used to help control for social and geographical dialect variation. Participants engaged in five tasks designed to compare the speech patterns of university-aged women and their mothers in different registers. A difference is hypothesized to be found in, both, the use of creaky voice cross-generationally, and between registers. Each participant read i. the Rainbow Passage, ii. a set of Harvard Sentences, and iii. a word list. These tasks were designed to provide an idea of the distribution of creaky voice in a formal discourse situation. Tasks iv. and v. are conversation tasks consisting of: a spot-the-differences picture task, and a route finding map task. These conversation tasks simulate a less formal discourse context. Annotations of the recordings were made which marked both the syllabic context in which creaky voice was produced and the length of time it was sustained at each occurrence. Using these annotations, global measurements of the usage of creaky voice were taken for each participant and compared across generations, registers and phonetic environments.

1. Introduction

This paper investigates the use of creaky voice by women in contemporary English. Typically, creaky voice quality has been associated with the diagnosis of voice disorders by speech language pathologists, yet recent studies have shown its increased use as part of an entire vocal range available to any speaker (Wolk *et al*. 2011, Gottliebson *et al*. 2006). In fact, creaky voice is such a normal voicing type that, in some tone languages (i.e., Hausa), it is used as a distinguishing feature between sounds (Ladefoged *et al*. 2010). Therefore, considering its popular manifestation among speakers of languages which do not use it to distinguish between phonemic categories, such as English, its use as a diagnostic tool among speech pathologists may be inappropriate.

Eckert (2004) conducted a socio-linguistic study investigating the use of certain characteristics present in the conversations of adolescents, such as the lexical item 'like' coupled with rising intonation, and the syntactic constructions 'I'm like...' and 'I'm all...'. It was shown that these items and constructions are "not just a random insertion," but a systematic addition that "serves to help organize the discourse" (Eckert 2004:7). Eckert (2004:6) also explains that these neologisms are tied to social identity. Not only is it interesting to consider the possibility that an increase in the popular usage of creaky voice can be likened to the use of the aforementioned constructions, it is important. Just as Eckert's (2004) study showed the correlation between identity and specific syntactic constructions, voice quality has likewise been correlated with socio-linguistic tendencies related to class (Laver 1980, Esling 1978, Trudgill 1974). Esling (1978) found that creaky voice was used more prevalently among those with higher social status in Edinburgh as opposed to the whispery or harsh voicing used among those with lower status. Beyond being a social marker, Laver (1980:1) describes an individual's voice as "an audible index of his identity, personality and mood." Divorcing speech acts from the discourse situations in which they occur, or the intonation and voice qualities with which they are produced, robs the researcher of a myriad of information that is contained within these extra-linguistic cues. However, these so-called extra-linguistic factors can be subtle hints at deeper issues and insights into the way we function as human beings within society. Despite the evidence for the importance of voice quality to one's identity, both socially and individually, very little work has been done on the use of different voice qualities used in dialogue.

This study looks at the usage of a specific voice quality, creaky voice, cross-generationally by mother-daughter pairs. The primary purpose of this study is to ascertain whether there is a difference in the usage of creaky voice both (i) between generations and (ii) between registers. The term *register* is generally used to refer to a variety of language that is interlocutor and context-dependent, such as that used in an informal discourse context versus a formal discourse context. For example, a student may use slang with other students, but then choose a less vernacular vocabulary when speaking to a professor (Platt & Platt 1975, Gregory & Carroll 1978). This is the definition assumed for this study. The secondary purpose of this study is to test whether there is a specific phonetic environment in which creaky voice is more likely to occur. For example, creaky voice is expected to be found in vowel articulation, but is it more likely to occur with liquids and glides than with nasals, or vice versa, and in which syllable position?

The remainder of this section serves to achieve three goals: To give a brief overview of the anatomy of the larynx; to outline the mechanics of voicing; and to describe the different voice qualities with which this study is concerned.

1.1 The Mechanics of Voicing

The airflow expelled from the lungs is essential to phonation. The pressure with which air is expelled from the lungs, in combination with the position of the vocal folds, also affects the manner in which the vocal folds vibrate and, thus, the resulting voice quality. There are a number of theories about vocal fold vibration such as: vibrating string theory, neurochonaxic theory, aerodynamic theory, myoelastic theory, muco-viscose, and flow-separation theories (Reetz & Jongman 2009).

A cycle of phonation, according to the aerodynamic and myoelastic theories, can be explained as follows: The first step in a single cycle of voicing is for the lateral cricoarytenoid muscles to tense causing the arytenoids to tilt down and inward, positioning the vocal folds for phonation (see Reetz & Jongman 2009: chapter 5 for an overview of the aforementioned theories). Airflow from the lungs forces the lower end of the vocal folds to open first and then, when the upper end of the vocal folds open, the *Bernoulli effect* kicks in. The Bernoulli effect emerges when a stream of particles flows through a narrow constriction. Within the constriction, the velocity of the air increases, which causes a drop in air pressure. This is important in phonation because the decrease in air pressure within the vocal folds, which form the constriction, creates a suction effect which pulls the vocal folds back together again. It is at this time, when the vocal folds come together, that the acoustic magic of phonation occurs (Reetz & Jongman 2009, Laver 1980). When the lower end of the folds are fully adducted, the upper end quicky follows suit. This closure allows for a build up of sub-glottal pressure and the cycle repeats itself (Reetz & Jongman 2009, Laver 1980).

It is this process, involving the position of the vocal folds and the air stream from the lungs, which allows for the occurrence of phonation. So then, what is the difference between a baby's cry and the singing of an aria? The answer lies in the setting of parameters. As mentioned above, when one of these parameters changes, the result is a change in phonation, or voice quality.

1.2 Voice Qualities and Their Characteristics

The are many voice qualities and a number of factors contribute to the differences in their production. These factors include: i. *sub-glottal pressure*, ii. *medial compression,* iii. *adductive tension,* and iv. *longitudinal tension*. This sub-section begins by defining these factors and then briefly describes the differences in pressure, compression and tension that are characteristic of three distinct voice qualities: breathy, modal and creaky.

As mentioned in the previous section, sub-glottal pressure, factor one, refers to the air pressure below the vocal folds in the sub-glottal system (the lungs). Medial compression, factor two, describes "the compressional pressure on the vocal processes of the arytenoid cartilages achieved by constriction of the lateral cricoarytenoid muscles and reinforced by tension in the lateral parts of the thyroarytenoid muscles" (Laver 1980:108). In other words, medial compression refers to the how tightly the vocal folds are pressed together. The vocal folds themselves have some form of medial compression inherent in

their musculature. Medial compression can be adjusted through the tensing of the thyroarytenoids, which move the arytenoid cartilages toward the thyroid, and the lateral cricoarytenoids, which cause adduction of the vocal folds (Laver 1980). Factor three, adductive tension, refers to how tightly the arytenoid cartilages are pressed together. Though the action of pressing the arytenoid cartilages together does bring the posterior end of the vocal folds together, adductive tension should not be confused with medial compression. This distinction is important because the arytenoid cartilages can remain open even when there is high medial compression on the vocal folds. For this reason, the section of the vocal folds attached to, and adducted by the arytenoid cartilages can be referred to as the *cartilaginous glottis* and the length of the folds, that run from the arytenoid cartilages to the thyroid cartilage can be referred to as the *ligamental glottis* (Laver 1980:107-108). Together, they make up the *full glottis* (Laver 1980:110). Adductive tension is increased by tensing the lateral cricoarytenoids and the transverse arytenoid muscles (Laver 1980). Longitudinal tension, factor four, is considered high when the vocal folds are stretched and low when they are relatively slack. The main factors in determining longitudinal tension are the vocalis muscles, the cricoid and thyroid cartilages, and the cricothyroid muscles (Laver 1980).

Modal voice, sometimes referred to as a "neutral mode of phonation" is characterized by regular vibration along all or most of the vocal folds (Laver 1980:110). There is low longitudinal tension, meaning the folds are shorter and thicker for the production of this type of phonation, and the other three factors, adductive tension, medial compression, and airflow, are all moderate. An increase in longitudinal tension in modal voice corresponds to an increase in pitch.

Breathy voice is produced with partial adduction along most or all of the length of the vocal folds (Reetz & Jongman 2009, Laver 1980). This means that adductive tension and medial compression are both low for this phonation type. Breathy voice, as its name suggests, has high airflow and the longitudinal tension can vary to adjust the pitch.

Creaky voice is characterized by irregular vibration of the vocal folds and occurs at the lower end of the F0 range. The irregularity in the vibration is caused by a combination of low sub-glottal pressure, high adductive tension along the cartilaginous glottis, and low longitudinal tension at the anterior end of the folds with high medial compression along the ligamental glottis (Ladefoged & Johnson 2010, Reetz & Jongman 2009, Laver 1980).

2. Methodology

As mentioned in the previous section, this study asks three questions: i. is there a cross-generational difference in the use of creaky voice among women; ii. is there a register difference in the use of creaky voice among women; and, iii. is there a phonetic environment in which creaky voice is more likely to occur? To address these questions, five mother-daughter pairs were audio-recorded while performing a series of reading and conversation tasks. The difference in task (reading versus conversation) is meant to represent a register change – formal versus informal, respectively. Cross-generational and cross-register differences are both expected. Based on pilot data and researcher observations, with respect to research question i., it is hypothesized that daughters will produce more creaky voice than their mothers. Regarding research question ii., it is

expected that both generations will produce more creaky voice in the informal discourse context.[3] Question iii. is being explored for information purposes.

2.1 Participants

Participants for this study were female students from the University of Calgary and their mothers. A total of five mother-daughter pairs were used in this study. Each participant was paid $20 for their participation. One mother reported a mild stutter, but no other hearing or speech impairments were reported. The reported stutter did not hinder the participant's production during any of the tasks. Ages ranged from 50 - 60 for mothers, with a mean age of 55, and 18 - 36 for daughters, with a mean age of 26. Four of the five daughters were from Alberta originally. The fifth daughter was originally from Ontario. Two of the mothers were from Alberta, two from Ontario and one from Michigan. The reported minimum length of time any one participant had lived in Calgary was four years. None of the participants were smokers. Four of the mother-daughter pairs were biologically related and one was adoptive. The purpose of choosing mother-daughter pairs for this study was to help control for dialect and socio-economic differences. This also allowed for the most direct comparison across generations.

2.2 Materials

Five tasks were used in this study – three reading tasks and two conversation tasks. The reading tasks were meant to simulate a formal discourse environment and the conversation tasks were meant to simulate an informal discourse environment. Each participant read the Rainbow Passage,[4] a set of Harvard Sentences and a word list and then participated in a picture task (spot the differences) and a map task. For this study, the third set of Harvard Sentences was chosen at random. Both the Rainbow Passage and the Harvard Sentences are phonetically balanced standard readings which are designed to test the production of connected speech. These readings are used in a number of production and comprehension tests such as speech evaluations, studying accents, speech exercises and testing language recognition software. The word list was a set of 44 monosyllabic words with no consonant clusters, such as: *rhyme* and *yak*. The word list was compiled to further test whether there is a phonetic environment in which creaky voice is more likely to occur.

2.3 Procedures

All tasks were performed in a sound attenuated booth with the experimenter present, so as to monitor the decibel (dB) level of the recordings. The reading tasks were performed separately, with only one participant present in the booth at a time, and the conversation tasks were performed with both members of the mother-daughter pairs. That is to say, the conversation occurred between the mother-daughter pairs, and not the participants and the experimenter. For the reading tasks, participants sat facing a Mac computer screen which displayed the reading tasks using a timed PowerPoint presentation. Participants spoke into a microphone which was mounted on a stand with a pop filter in front of the microphone. After reading the Rainbow Passage, participants pressed the *enter* key once on

[3] These phenomena have been part of popular discussion, but have not yet made it into the literature.
[4] See Appendix A-E for more information about the materials.

the Mac keyboard to advance to the next slide and to start the timed presentation of the Harvard Sentences and the word list. The PowerPoint slides were set to advance at 5 second intervals for the Harvard Sentences and 3 second intervals for the word list. A timed presentation was used to help minimize list intonation and background noise, which was exhibited with the use of paper copies of the material in the pilot study.

After both participants had separately completed the reading tasks, they were both asked to enter the booth for their participation in the conversation tasks. The mother-daughter pairs sat facing each other and spoke directly into their own designated Shure SM-48 microphones, which were mounted on stands with pop filters in front of them. The microphones fed into a Computerized Speech Lab (CSL) model 4500 box for analog-to-digital conversion. Conversations were recorded in stereo using Adobe Audition and saved as .wav files for analysis.

For the picture task, each participant was presented with an image that varied in 10 different aspects. Participants were asked to use verbal skills only to locate the differences in the pictures and not to look at each other's image. They were asked to locate five of the 10 differences before finishing the task as some of the differences were too subtle to find in this manner. The purpose of not allowing the participants to see one another's image was to insure that conversation would be used to perform the task in lieu of pointing and the use of deictics, which can minimize the amount of conversation used.

For the map task, participants were given the same map, one with a route and one without. The person who received the map with the route was asked to give the other participant directions from point A to point B. Each participant took a turn being the 'navigator' with a different map. Again, participants were asked not to look at each other's image, but to use verbal skills to complete the task.

2.4 Analysis

The data collected were analyzed using Praat (Boersma & Weenink 2010). Annotations of all sound files were made using the following tiers: words, creaky, breathy, and modal. This was done so as to transcribe when each voice quality occurred. Since this study focuses on the usage of creaky voice, the creaky tier encoded further syllabic information such as onset (O), nucleus (N), coda (C), syllable boundary (.), and word boundary (#) to mark where this phonation type was occurring within the word. For example, the word 'phonation' has three syllables which orthographically correspond to '#pho.na.tion#'. Figure 1 below shows an example of the annotation used in this study.

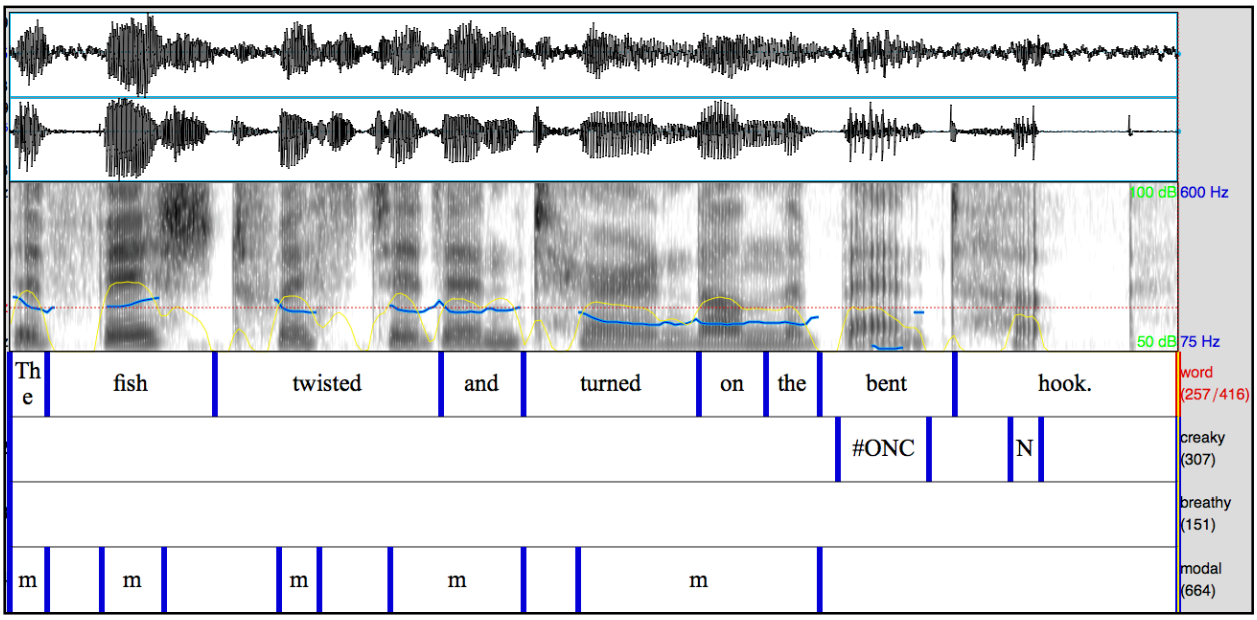


Figure 1: Notation example of Harvard Sentence, Set no. 3, sentence no. 2. This figure shows the annotation of the four tiers (words, creaky, breathy, and modal), the syllable position and phonation type.

For ease of analysis, any voice quality observed in the data that fell outside of one of the previously described voice qualities (creaky, breathy, modal), was grouped under either modal or breathy. For example, tense (or pressed) voice, any cracks, squeaks or inconsistencies that were clearly not creaky were classified as modal and marked with an '*m*' inside the tier. Phrase final devoicing was classified as breathy with a '*b*' inside the tier.[5]

A script was run that took global measurements of each voice quality used. That is to say, since all voiced segments were marked as one of the previously mentioned voice qualities (creaky, breathy, modal), this script measured the percentage of all voicing that was creaky, breathy, or modal. This allowed for the cross-generational and register comparison of the use of creaky voice. This same script also compiled statistics on the syllable position and segment type in which creaky voice occurred. This allowed for the analysis of the phonetic environment in which creaky voice was produced.

1. Results

3.1 Cross-Generational & Cross-Register Data

The global measurements of voice quality gave total percentages of each phonation type used during the tasks. Though the sample size was not large enough to run a sufficiently powerful statistical analysis, findings from the voice quality analysis revealed that the daughters produced 7% more creaky voice than the mothers overall. A slight register difference was found for the mothers' data in which the participants produced 2% more creaky voice overall during the conversation tasks (informal register). The data for the daughters shows a 4% cross-register difference. See Figures 2 and 3:

[5] See Appendix F.

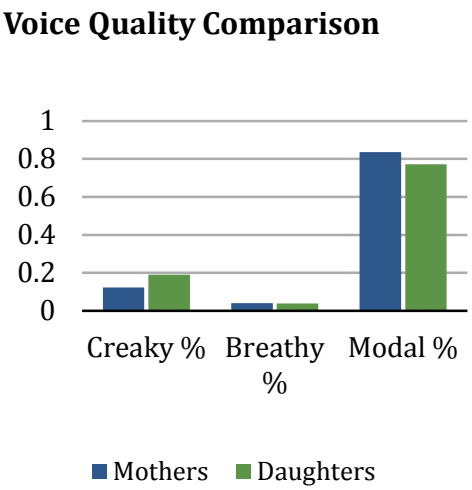


Figure 2: A voice quality comparison which shows the percentage of voice qualities used in all tasks by all participants. Mothers: 12%, 4% and 84%. Daughters: Mothers:19%, 4%, 77%.

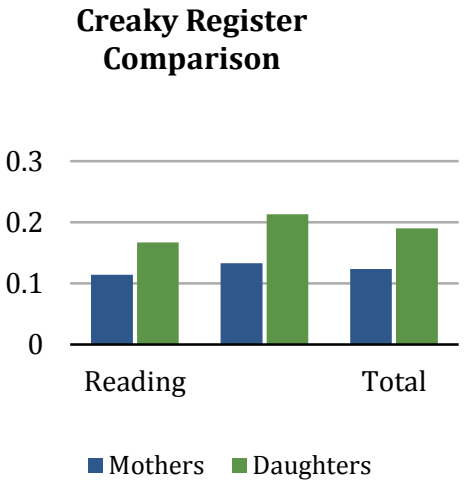


Figure 3: A register comparison between tasks which shows the percentage of creaky voice used in each task by all mothers and daughters. 11%, 13%, 12%. Daughters: 17%, 21%, 19%.

Figure 3 also indicates that daughters produced 6% more creaky voice in reading tasks (formal register) and 8% more in conversation tasks (informal register) than the mothers.

Within each of the mother-daughter pairs, there was quite a lot of variation in the production of creaky voice. Figures 4-6 summarize this data:

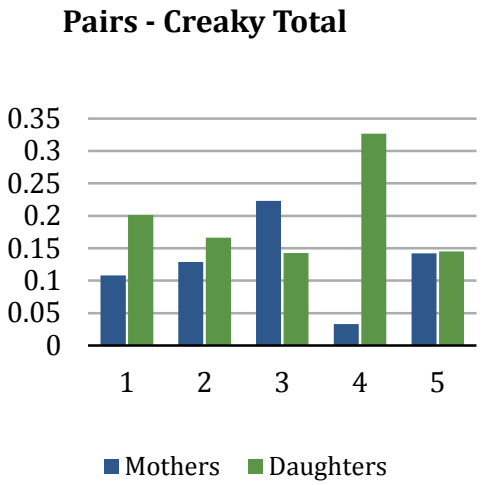


Figure 4: A mother-daughter pair comparison which shows the percentage of creaky voice used in all tasks by each mother and daughter. (1): 11%, 20%. (2): 13%, 17%. (3): 22%, 12%. (4): 3%, 32%. (5): 14%, 15%.

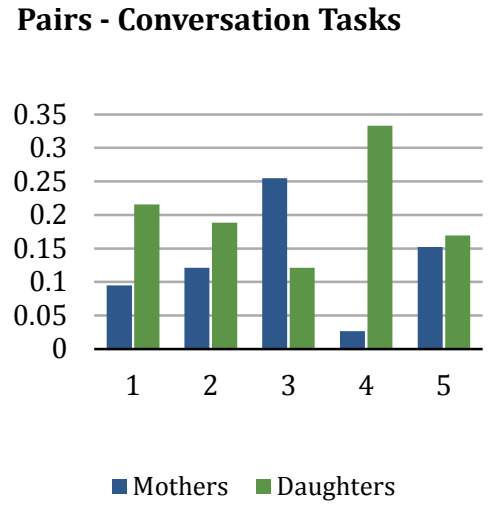


Figure 5: A mother-daughter pair comparison of conversation tasks which shows the percentage of creaky voice used in conversation tasks by each mother and daughter. (1): 10%, 22%. (2): 12%, 19%. (3): 25%, 12%. (4): 3%, 33%. (5): 15%, 7%.

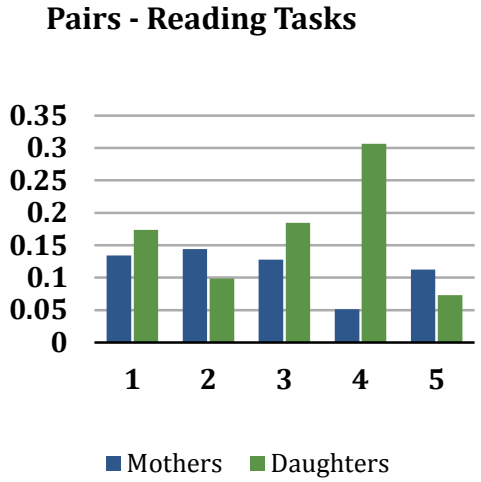


Figure 6: A mother-daughter pair comparison of reading tasks which shows the percentage of creaky voice used in reading tasks by each mother and daughter. (1): 13%, 17%. (2): 14%, 10%. (3): 12%, 18%. (4): 5%, 31%. (5): 11%, 7%.

It is interesting to note that the mothers in mother-daughter pairs 2 and 5 exhibited more creaky phonation during the reading tasks (formal discourse context) then their daughters. Also, the mother in mother-daughter pair 3 produced double the amount of creaky phonation during the conversation tasks (informal discourse context). A couple of possible points of interest from the participant questionnaires are listed here[6]: the mother from pair 3, who produced far more creaky voice than her daughter during the conversation tasks, was from Michigan; pair 4, which exhibited the greatest cross-generational difference in the use of creaky voice, was the oldest pair of the subject pool (the daughter was 36 and the mother was 60 years of age); and the mother from pair 5, which exhibited almost equal amounts of creaky voice cross-generationally, was the only mother who spoke two languages (English and French).

3.2 Syllable Position & Segmental Data

The number of occurrences of each syllable position was totaled (onset, nucleus, and coda) and then compared with the number of occurrences in which creaky voice was produced in each syllable position. Results for syllable position revealed a 6% increase in the use of creaky voice from onset position to coda position for both mothers and daughters with a 3% cross-generational difference between mothers and daughters in both onset and coda position. See Figure 7:

[6] See Appendix G.

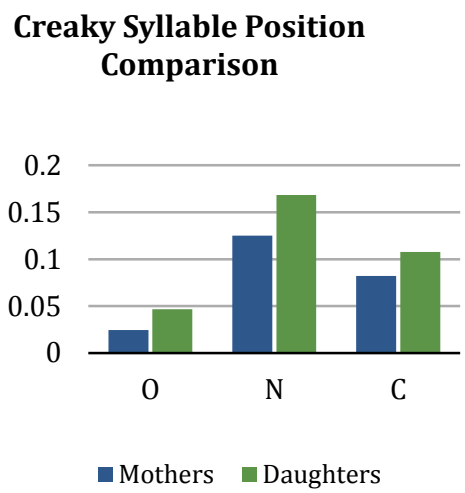


Figure 7: A syllable position comparison which shows the percentage of creaky voice used in onset (O), nucleus (N) and coda (C) positions by both mothers and daughters. (O): 2%, 5%. (N): 13%, 17%. (C): 8%, 11%.

Figures 8-12 summarize the segment data which indicate where creaky voice was produced. With respect to Figure 8, *% Creaky Stops*, a stop was included in the creaky voice portion of the annotations if creakiness was clearly heard in the formant transitions leading into or out of the stop consonant. This provides information about the immediate environment in which creaky voice took place. So, for example, in Figure 8 it can be seen that both mothers and daughters produced creaky voice before or after a [t] more frequently than any other stop consonant (7% for mothers and 8% for daughters).

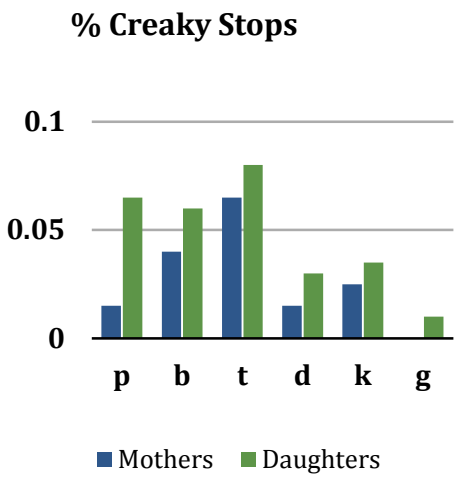


Figure 8: A segment comparison which shows the percentage of creaky voice used immediately before before or after a stop consonant. [p]: 2%, 7%. [b]: 4%, 6%. [t]: 7%, 8%. [d]: 2%, 3%. [k]: 3%, 4%. [g]: 0%, 1%.

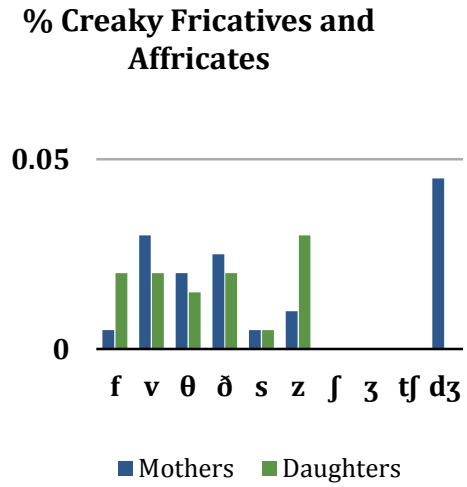


Figure 9: A segment comparison which shows the percentage of creaky voice used immediately or after fricatives and affricates. [f]: 1%, 2%. [v]: 3%, 2%. [θ]: 2%, 2%. [ð]: 3%, 2%. [s]: 1%, 1%. [z]: 1%, 3%. [ʃ]: 0%, 0%. [ʒ]: 0%, 0%. [tʃ]: 0%, 0%. [dʒ]: 5%, 0%.

Creaky voice was produced more frequently immediately before or after [p,b,t] by daughters (an average of 7% of the time) than [d, k, g] (an average of 2% of the time). Mothers produced creaky voice an average of 4% of the time before or after [p, b, t] and an average of 1% of the time before or after [d, k, g]. Both mothers and daughters produced creaky voice more frequently in the immediate environment of [t], 7% and 8% respectively. Both mothers and daughters produced the least amount of creaky voice in the immediate environment of [g], 0% and 1% respectively.

Figure 9 shows the percentage of creaky voice produced in the immediate environment of fricatives and affricates. The data shows a steady low usage of creaky voice in these environments for both mothers and daughters which range from 0-3% for fricatives. Neither mothers nor daughters produced creaky voice immediately before or after [ʃ] or [ʒ]. As Figure 9 indicates, daughters did not seem to use affricates as a 'creakable' environment, while mothers produced creaky voice only immediately before or after the voiced affricate [dʒ] 5% of the time.

Figure 10 shows the use of creaky voice during nasal consonants. Both mothers and daughters produced the most creak while articulating the alveolar [n], 11% and 14% respectively. Daughters produced creak 13% of the time with both [m] and [ŋ], while mothers produced the least amount of creak with [m] at 5%.

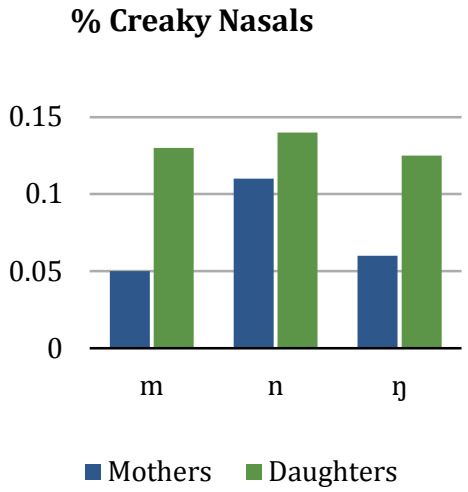


Figure 10 (left): A segment comparison which shows the percentage of creaky voice used during the articulation of nasal consonants. [m]: 5%, 13%. [n]: 11%, 14%. [ŋ]: 6%, 13%.

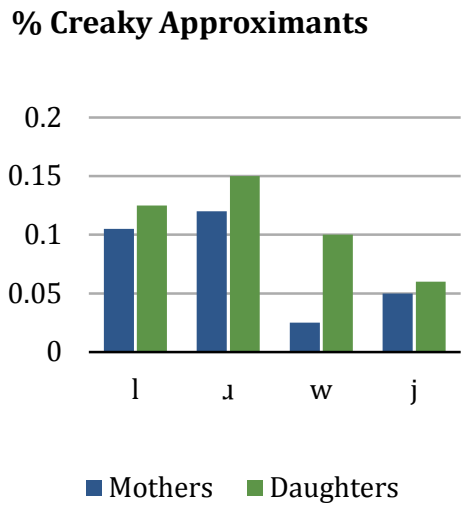


Figure 11 (right): A segment comparison which shows the percentage of creaky voice used during the articulation of approximants. [l]: 11%, 13%. [ɹ]: 12%, 15%. [w]: 3%, 10%. [j]: 5%, 6%.

Approximants are considered in Figure 11. The data shows a greater use of creaky voice during articulation of liquids for both mothers and daughters than for glides with [ɹ] showing the highest usage of creaky voice at 12% for mothers and 15% for daughters.

A comparison of low and high vowels is shown in Figure 12. Mothers produced creaky voice 6% more frequently in low vowels than in high vowels. Daughters produced creaky voice 9% more frequently in low vowels than in high vowels.

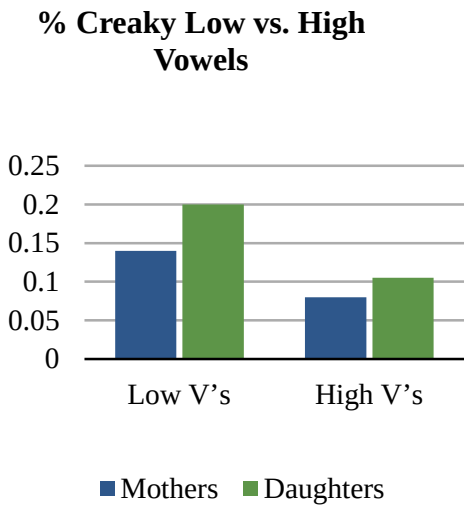


Figure 12 (left) : A segment comparison which shows the percentage of creaky voice used during the articulation of low vs. high vowels. Low V's: 14%, 20%. High V's: 8%, 11%.

4. Discussion

The data indicates support for the two hypotheses that i. university-aged women produce more creaky voice than their mothers and ii. women produce more creaky voice in an informal discourse environment than in a formal discourse environment. Here, it was found that daughters produced creaky voice 7% more frequently than mothers overall with variation within each mother-daughter pair. The results presented in Figures 4-6, which show the mother-daughter pair data, indicates that the oldest pair (pair 4) showed the greatest cross-generational difference in the use of creaky voice. One thing to consider here is the possibility that further investigations into this topic may be better served by choosing an older subject pool. With respect to the anomaly in pair 3, where the mother produced double the amount of creaky voice in the conversation tasks (informal discourse context), it would be interesting to see if Americans are more likely to exhibit this phenomenon within an older cross-generational subject pool.

The syllable position and segment data show a tendency for creaky voice to be produced more often in coda position than in onset position. In terms of the nucleus, which is where creaky voice is expected to occur, it seems there is a 'preferred' natural class among vowels where creaky voice is more likely to occur. The data showed a 6% and 9% increase for mothers and daughters respectively in the production of creaky voice during low vowels compared to high vowels. Further investigation needs to be done on this matter, but I speculate that tongue position may have a large role to play in this difference (Honda 2004).

This factor may also explain the differences found in the data for approximants. [w] and [j] correspond to the high vowels [u] and [i] respectively, which may account for the lower rate with which creaky voice was produced during their articulation in comparison with the liquids [ɹ] and [l]. That is to say, [ɹ] and [l] have been found to exhibit a more similar tongue root position to that of low vowels than high vowels (Gick *et al.* 2002). More research needs to be done on this matter to ascertain the physiological correlation between

tongue root retraction and creaky voice. For instance, does the retraction of the tongue root put muscular demands on the larynx that induces this phonation type?

This same factor, tongue position, could also explain the differences with respect to the data for stops. The velar environment [k, g] exhibited the lowest percentage of creaky voice. The tongue position for these consonants could be likened to that of high vowels. However, this poses a potential problem for the nasal stop data in which [ŋ] had a high percentage of creaky voice production. Before and after [t] was found to be the most creaked environment for stops. In English, allophonic variation is found with alveolar and glottal stops, which may partially account for this phenomenon. For example, [t] can be produced as [ʔ] before a syllabic [n̩] as in the lexical item 'button' in RP (Received Pronunciation) (Ladefoged & Johnson 2010). It could also be argued that glottalization could co-occur with unreleased [tˀ] in word final position. So, for example, the word 'spot', when produced as [spɑtˀ] with an unreleased final [tˀ], could also be undergoing glottalization at the same time as the final stop closure as in [spɑtʔˀ]. If this is the case, this glottalization would account for the fact that [t] had the highest amount of creaky voice occurring before and after it. More investigation needs to be done here with respect to creaky phonation in the immediate environment of stop consonants.

5. Conclusion

There are many future directions in which this research could go. It would be interesting to see if there is a greater cross-generational difference among a slightly older subject pool. Further research could investigate the potential of this phenomenon being tied to Feminism and social identity. Female gender roles have changed and this change could not have occurred without a shift in the female identity (MacIvor 2003). If females are characteristically producing a different voice quality, one that is on the low end of their vocal range as creaky voice is, it could be a reflection of the female attempt to fit into a masculine society (MacIvor 2003). Further consideration can be made for cultures in which the Feminist movement has not taken hold to the same extent to which it has in first world countries. How would the use of creaky voice manifest itself among women of other languages whose cultures are primarily patriarchal?

As was evidenced by the data, a slight register difference was found. Future research could also investigate whether the amount of creaky phonation that is produced during discourse is interlocutor dependent. For instance, would a woman be more apt to increase her production of creaky voice when speaking to a male professor than a female one? What about a well dressed male stranger vs. a poorly dressed one? Would voice quality imitation or accommodation come into play more with an attractive conversation partner than with an unattractive one?

The mother in pair 5, who produced an almost equal amount of creaky voice as her daughter, was the only mother who spoke two languages. It would be interesting to investigate whether bilinguals exhibit more or less creaky voice than monolinguals. Or if there is a correlation between the amount of creaky voice produced and the native-language a woman speaks. For instance, would Francophones be more or less apt to produce creaky voice than Anglophones?

A phonetic environment in which creaky phonation was used more frequently was observed. This was found to be in coda position (vs. onset position), in the alveolar place of articulation for stops and nasals, in liquids (vs. glides) and in low vowels (vs. high vowels). A further look into the physiological reasons for the variation in creaky production among segments also needs to take place. Also, investigation into the glottalization of [t] could shed light on its effect on the voice quality of surrounding segments.

In sum, this study investigated the usage of creaky voice across generations and registers among women. Evidence was found to support the cross-generational variation hypothesis in which daughters produced more creaky voice than mothers. A slight register difference was also found in which creaky voice was produced more frequently in an informal discourse situation. This evidence supports the idea that voice quality is not merely an extra-linguistic factor that has little or no bearing on the way language is used. On the contrary, phonation patterns seem to be a very real part of the ebb and flow of human social interaction and communication.

References

Boberg, Charles. 2008. Regional phonetic differentiation in standard Canadian English. *Journal of English Linguistics*, 36(2), 129-154.

Curtin, Susanne & Stephen Winters. 2011. The perception of syntactic disambiguation in infant- directed speech. Talk presented at the Annual Conference of the Canadian Acoustical Association, Quebec City, Quebec. October 12.

Eckert, Penelope. 2004. Adolescent language. In E. Finegan and J. Rickford (eds.), *Language in the USA*. New York: Cambridge University Press.

Esling, John H. 1978. Voice quality in Edinburgh: A sociolinguistic and phonetic study, Ph.D. dissertation, University of Edinburgh.

Gottliebson, Renee O., Linda Lee, Barbara Weinrich & Jessica Sanders. 2006. Voice problems of future speech-language pathologists. *Journal of Voice,* 21(6), 699-704.

Gick, Bryan, A. Min Kang & Douglas H. Whalen. 2002. MRI evidence for commonality in the post-oral articulations of English vowels and liquids. *Journal of Phonetics*, 30, 357-371.

Gregory, Michael & Susanne Carroll. 1978. *Language and situation: Language varieties and their social contexts.* Boston, MA: Routledge.

Honda, Kiyoshi 2004. Physiological factors causing tonal characteristics of speech: From global to local prosody. *Processing Speech Prosody*. Nara, Japan.

Ladefoged, Peter & Keith Johnson. 2010. *A course in phonetics*. (6th ed.). Boston: Thomson Wadsworth.

Laver, John. 1980. *The phonetic description of voice quality.* Cambridge, UK: Cambridge University Press.

MacIvor, Heather. 2003. Women in Canada: Two steps forward, one step back. In K. Pryke & W. Soderland (eds.), *Profiles of Canada*. Toronto, ON: Canadian Scholars' Press. 145-177.

Platt, John T. & Heidi K. Platt. 1975. *The social significance of speech*. New York, NY: North-Holland.

Reetz, Henning & Allard Jongman. 2009. *Phonetics: Transcription, production, acoustics, and perception*. Malden: Wiley-Blackwell.

Simberg, Susanna, Anneli Laine, Eeva Sala, & Anna-Maija Ronnemaa. 2000. Prevalence of voice disorders among future teachers. *Journal of Voice,* 14(2), 231-235.

Simberg, Susanna, Eeva Sala, & Anna-Maija Ronnemaa. 2004. A comparison of the prevalence of vocal symptoms among teacher students and other university students. *Journal of Voice,* 18, 363-368.

Trudgill, Peter. 1974. *The social differentiation in Norwich*, Cambridge University Press.

Wolk, Lesley, Nassima B. Abdelli-Beruh, & Dianne Slavin. 2012. Habitual use of vocal fry in young adult female speakers. *Journal of Voice,* 26(3), 111-116.

Appendix A: Rainbow Passage

When the sunlight strikes raindrops in the air, they act like a prism and form a rainbow. The rainbow is a division of white light into many beautiful colors. These take the shape of a long, round arch, with its path high above and its two ends apparently beyond the horizon. There is, according to legend, a boiling pot of gold at one end. People look, but no one ever finds it. When a man looks for something beyond his reach, his friends say he is looking for the pot of gold at the end of the rainbow. Throughout the centuries men have explained the rainbow in various ways. Some have accepted it as a miracle without physical explanation. The Greeks used to imagine that it was a sign from the gods to foretell war or heavy rain. The Norsemen considered the rainbow as a bridge over which the gods passed from earth to their home in the sky. Other men have tried to explain the phenomenon physically. Aristotle thought that the rainbow was caused by reflection of the sun's rays by the rain. Since then, physicists have found that it is not reflection, but refraction by the raindrops, which causes the rainbow. Many complicated ideas about the rainbow have been formed. The difference in the rainbow depends considerably upon the size of the water drops, where the width of the colored band increases as the size of the drops increase. The actual primary rainbow observed is said to the effect of superposition of a number of bows. If the red of the second bow falls upon the green of the first, the results is to give a bow with abnormally wide yellow band, since red and green lights when mixed form yellow. This is a very common type of bow, one showing mainly red and yellow, with little or no green or blue.

Appendix B: Harvard Sentences (Set no. 3)

1. The small pup gnawed a hole in the sock.
2. The fish twisted and turned on the bent hook.
3. Press the pants and sew a button on the vest.
4. The swan dive was far short of perfect.
5. The beauty of the view stunned the young boy.
6. Two blue fish swam in the tank.
7. Her purse was full of useless trash.
8. The colt reared and threw the tall rider.
9. It snowed, rained, and hailed the same morning.
10. Read verse out loud for pleasure.

Appendix C: Word List

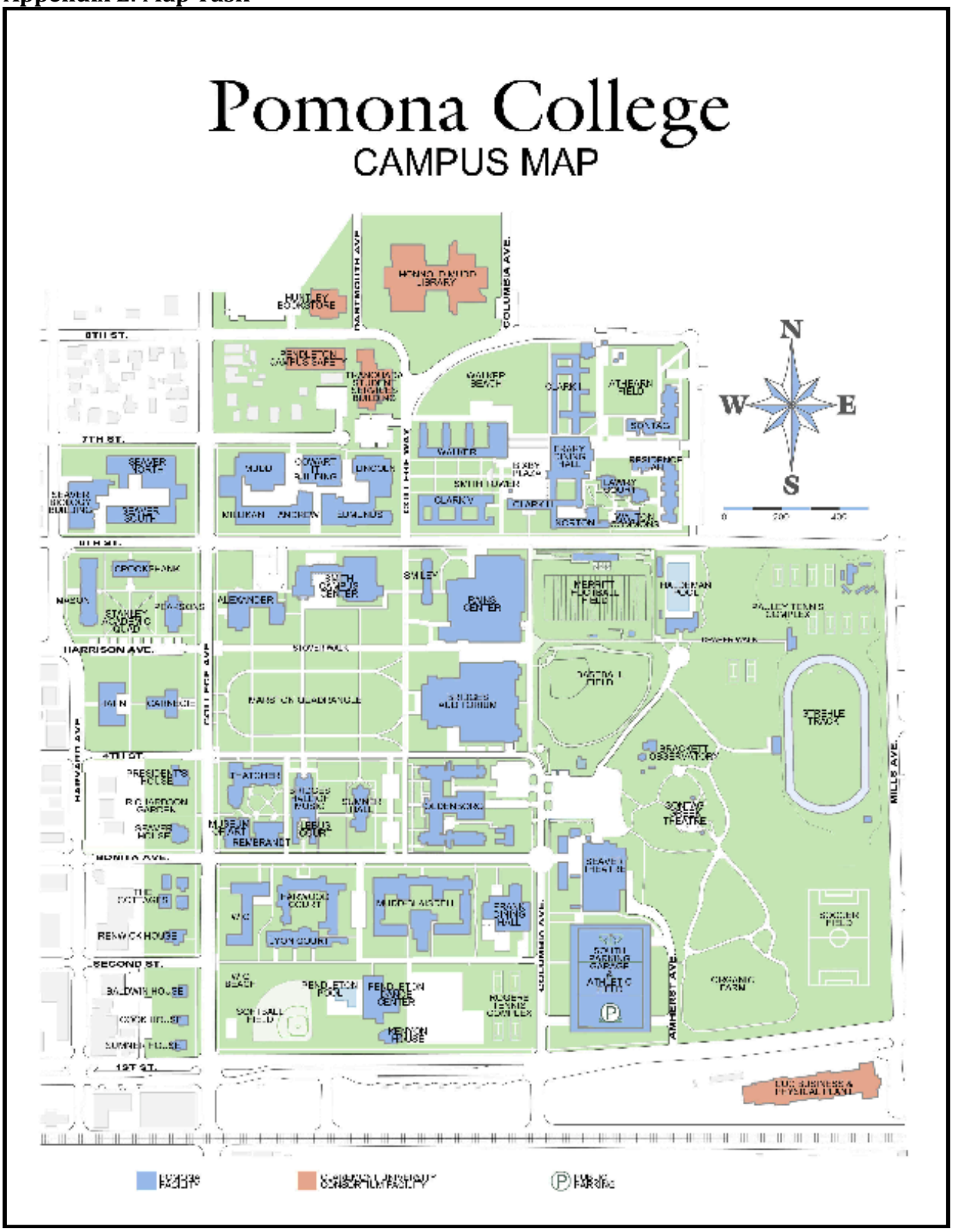
The following is the word list compiled. The only simple onset not used was the labio-velar glide [w].

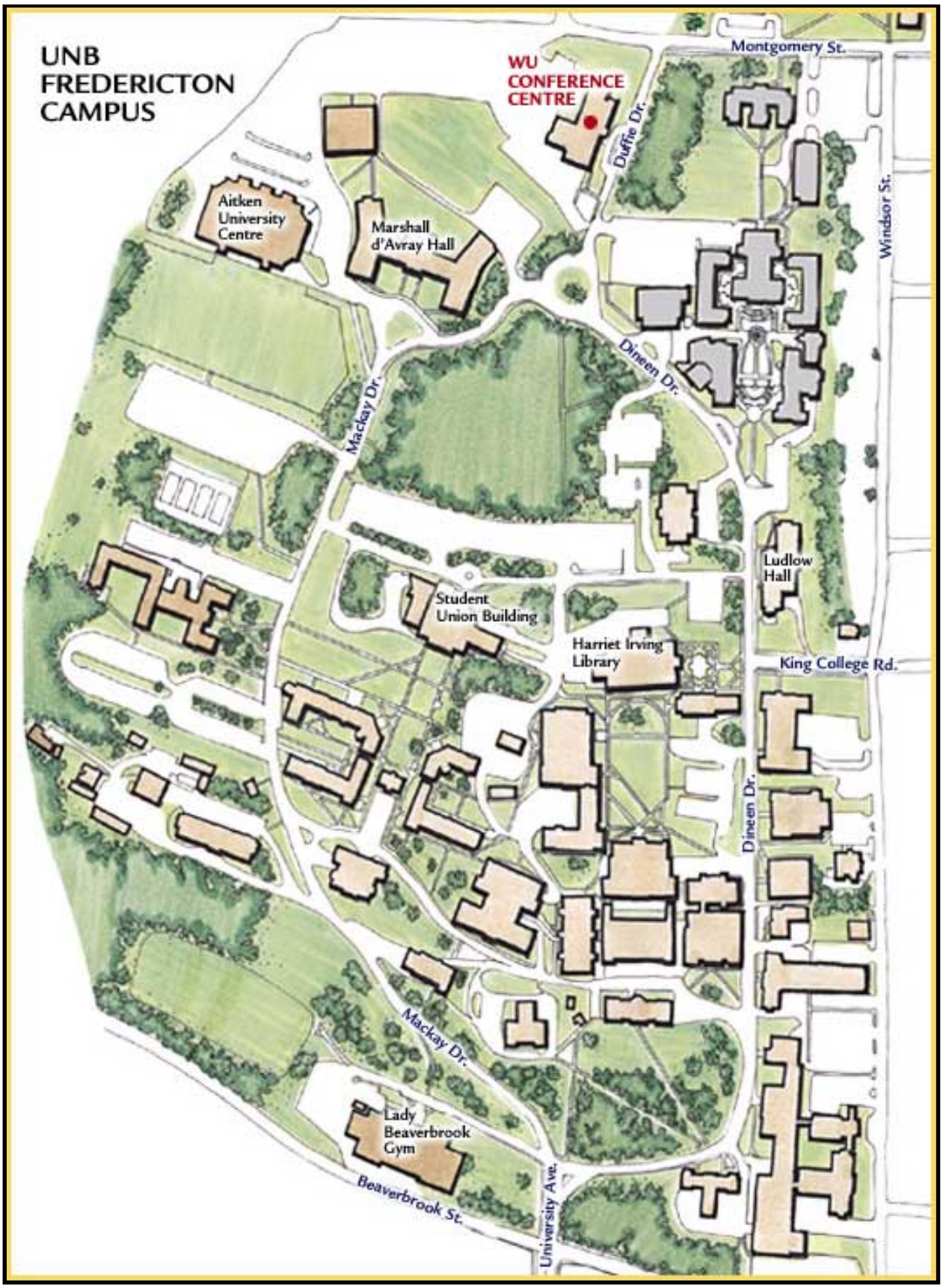
Onset	Lexeme	Nucleus
Stops		
[p]	pave	[eɪ]
	puck	[ʌ]
[b]	badge	[æ]
	boil	[oɪ]
[t]	tout	[aʊ]
	tone	[oʊ]
[d]	dowse	[oʊ]/[aʊ]
	dice	[aɪ]/[əɪ]
[k]	coin	[oɪ]
	core	[ɔ]
[g]	goose	[u]
	gull	[ʌ]
[ʔ] / null	own	[oʊ]
	out	[aʊ]/[əʊ]
Fricatives		
[f]	fair	[ɛ]
	foot	[ʊ]
[v]	void	[oɪ]
	vain	[eɪ]
[θ]	thumb	[ʌ]
	thin	[ɪ]
[ð]	these	[i]
	then	[ɛ]
[s]	sauce	[ɑ]
	sob	[ɑ]
[z]	zeal	[i]
	zip	[ɪ]
[ʃ]	should	[ʊ]
	shoot	[u]
[ʒ]	-	-
	-	-
[h]	hood	[ʊ]
	house	[aʊ]/[əʊ]
Affricates		
[tʃ]	cheek	[i]
	choke	[oʊ]
[dʒ]	gem	[ɛ]
	gym	[ɪ]
Nasals		
[m]	mauve	[oʊ]/[ɑ]

Onset	Lexeme	Nucleus
	mop	[ɑ]
[n]	need	[i]
	night	[aɪ]/[əɪ]
[ŋ]	-	-
	-	-
Glides		
[j]	yore	[ɔ]
	yak	[æ]
Liquids		
[l]	lore	[ɔ]
	lush	[ʌ]
[ɹ]	rot	[ɑ]
	rhyme	[aɪ]

Appendix D: Picture Task

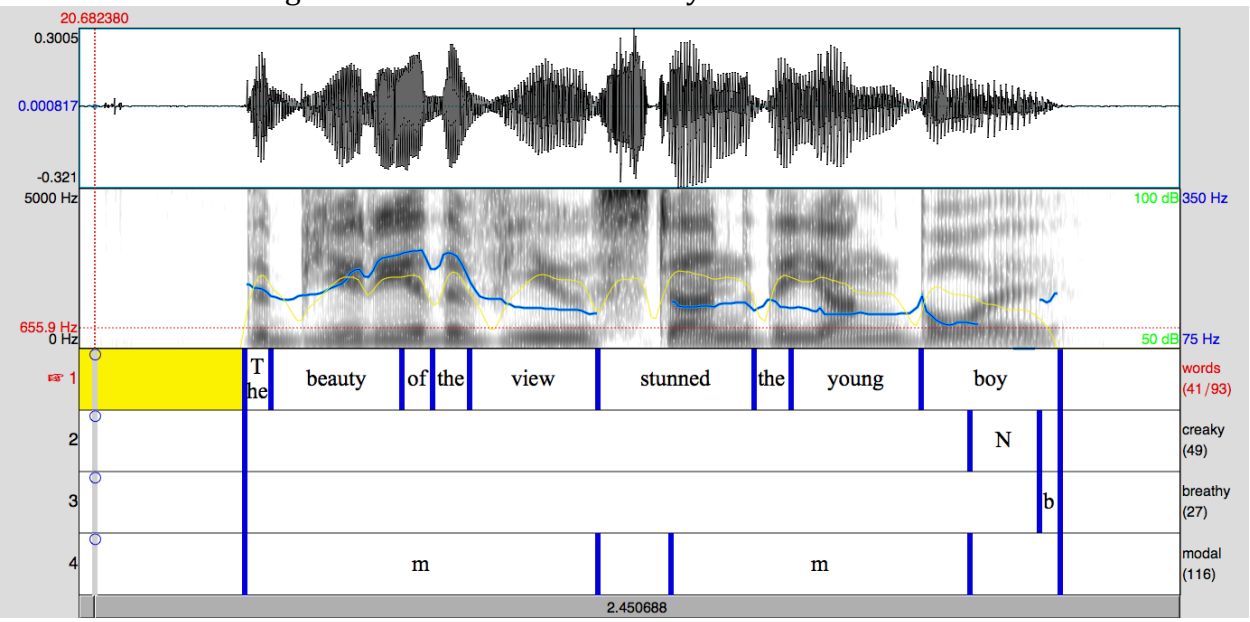
Appendix E: Map Task



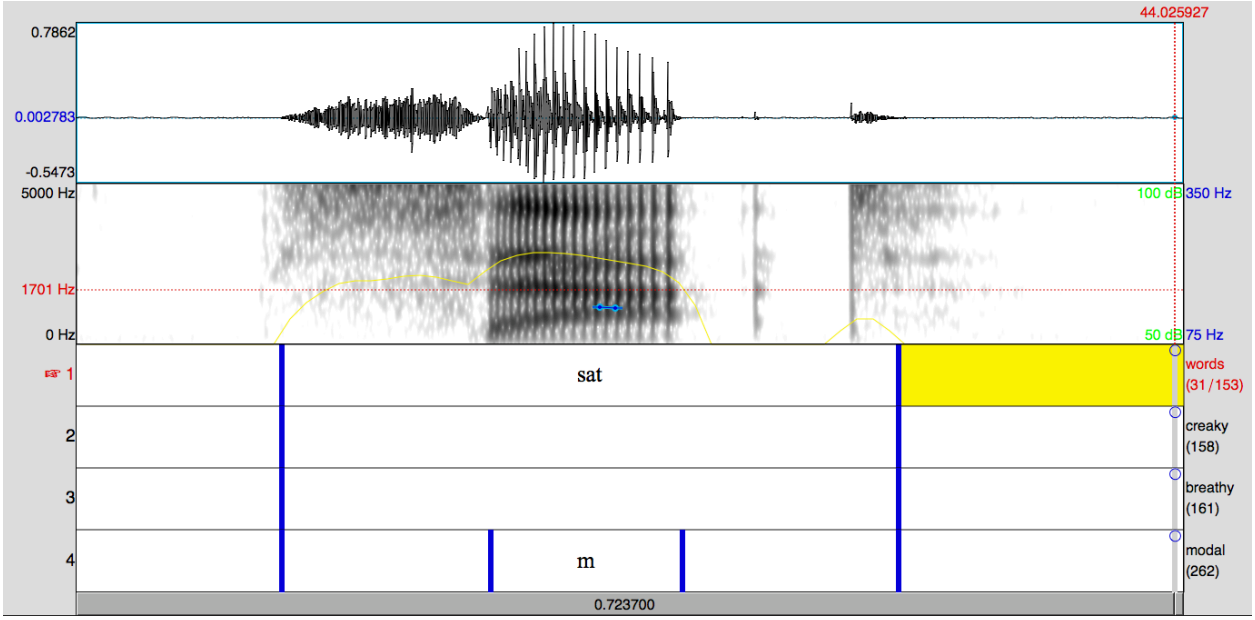

UNB
FREDERICTON
CAMPUS
WU
CONFERENCE
CENTRE
Montgomery St.
Duffie Dr.
Windsor St.
Aitken
University
Centre
Marshall
d'Avray Hall
Dineen Dr.
Mackay Dr.
Ludlow
Hall
Student
Union Building
Harriet Irving
Library
King College Rd.
Dineen Dr.
Mackay Dr.
Lady
Beaverbrook
Gym
Beaverbrook St.
University Ave.

Appendix F: Other Phonation Types

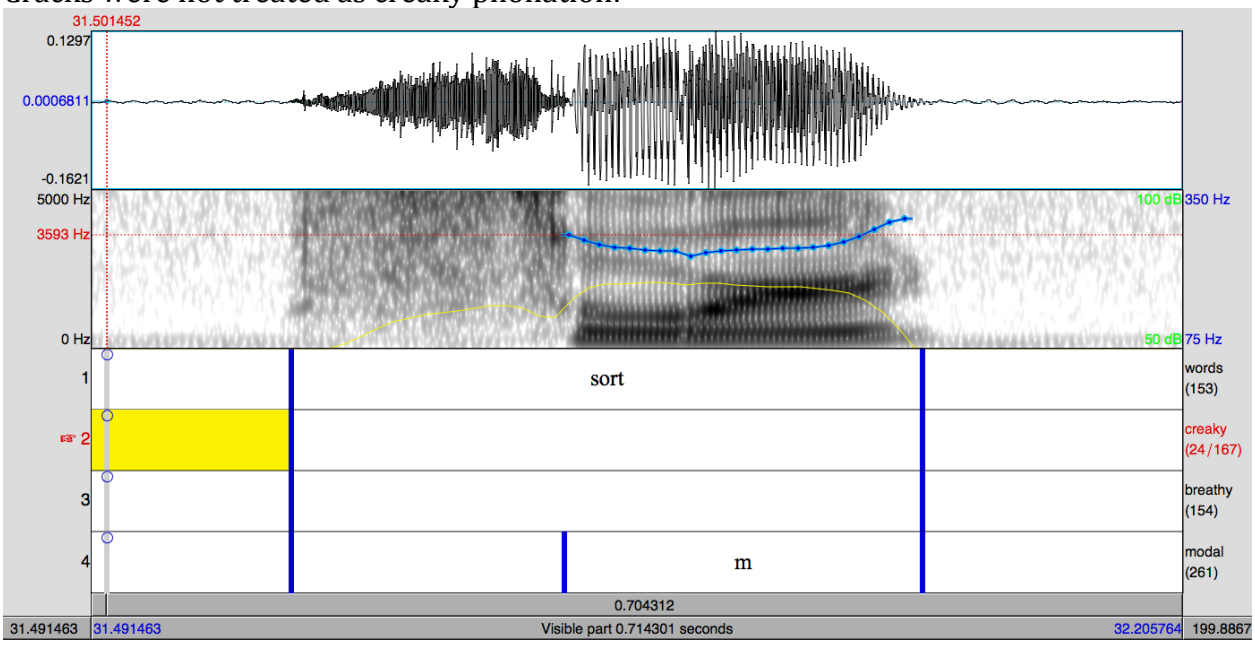
Phrase final devoicing was marked as *b* for 'breathy':



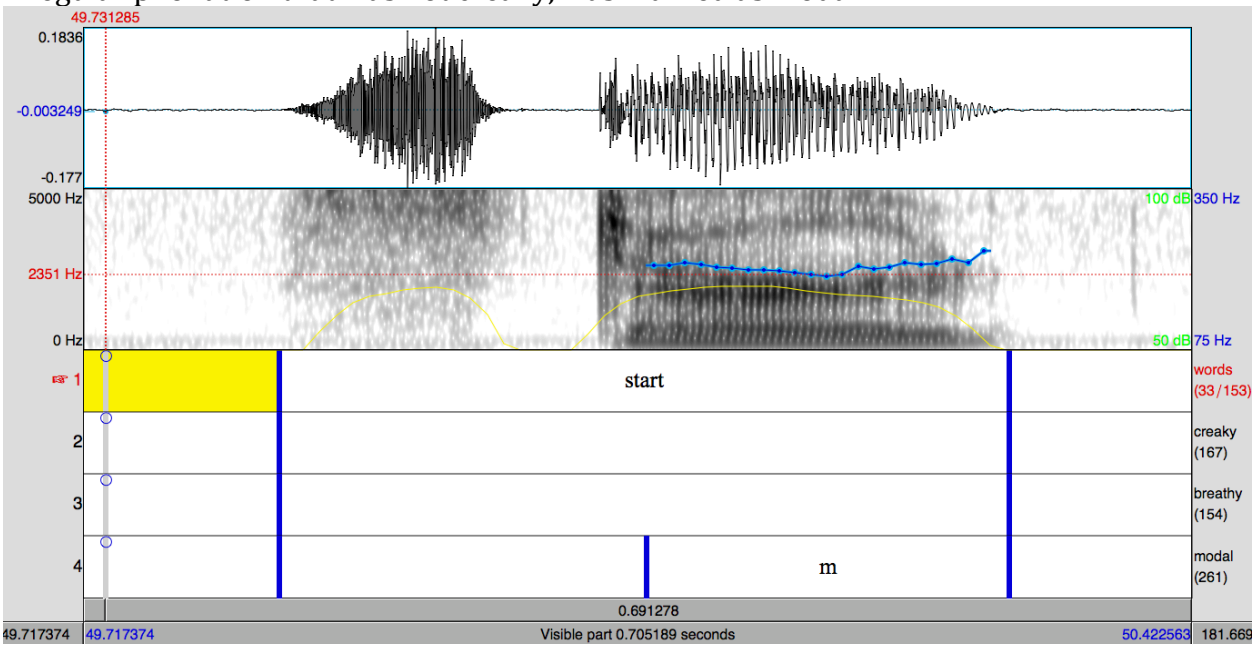
Pressed voice was labelled as *m* for 'modal':



Cracks were not treated as creaky phonation:



Irregular phonation that was not creaky, was marked as modal:



Appendix G: Participant Questionnaire Data

D1 = Daughter of mother-daughter pair 1, D2 = Daughter of mother-daughter pair 2...
M1 = Mother of mother-daughter pair 1, M2 = Mother of mother-daughter pair 2...

	D1	**D2**	**D3**	**D4**	**D5**
Age	25	23	18	36	30
From	Calgary	Calgary	Windsor	Edmonton	Calgary
Calgary	25	23	5	16	20
Languages	English, French	English	English	English, French	English
Smoker	No	No	No	No	No
Smoked	N/A	N/A	N/A	N/A	N/A
Occupation	Law Student	Student	Student	Graduate Student	Business Analyst
Education	BA - IR	High School	High School	BA - Honours French	Career College - Office Administration
Musical	flute, guitar	No	No	No	No
Speech/Hearing Impairments	No	No	No	No	No
Biological/Adoptive	Biological	Biological	Biological	Adoptive	Biological
	M1	**M2**	**M3**	**M4**	**M5**
Age	55	54	54	60	50
From	Ontario	Calgary	Detroit	Edmonton	Ottawa
Calgary	27	54	4	55	20
Languages	English	English	English	English	English, French
Smoker	No	No	No	No	No
Smoked	N/A	N/A	N/A	N/A	N/A
Occupation	Retired Farmer	Teacher	Admin	Teacher	HR
Education	MEd - Reading & Language	MEd	Associates Degree	BEd - Elem. Physical Ed, Dip. ECE	BA - Admin., Organizations
Musical	No	No	No	No	flute
Speech/Hearing Impairments	No	No	No	No	Yes - mild stutter
Biological/Adoptive	Biological	Biological	Biological	Adoptive	Biological

Sarah Greer
greer.sarah@gmail.com

Stephen J. Winters
swinters@ucalgary.ca

University of Calgary
SS 820 - Department of Linguistics, Languages and Cultures
2500 University Dr. NW
Calgary, Alberta
T2N 1N4
Canada