CREAK IN THE RAIN:

PHONATION IN OREGON ENGLISH

JOHN M. RIEBOLD

A MAJOR RESEARCH PAPER SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

GRADUATE PROGRAM IN THEORETICAL LINGUISTICS

YORK UNIVERSITY,

TORONTO, ONTARIO

AUGUST 25TH, 2009

SUPERVISED BY PHILIPP ANGERMEYER, ASSISTANT PROFESSOR OF LINGUISTICS, DEPARTMENT OF LANGUAGES, LITERATURES AND LINGUISTICS

ABSTRACT

Although English in the US Pacific Northwest has seen comparatively little research, many scholars in the area have mentioned local speakers' frequent use of creaky voice (Ingle, Wright, & Wassink 2005, J. Conn, personal communication, February 9, 2009, Ward 2003). Despite findings showing that creaky voice and other phonations are significant in other languages and dialects of English (cf. Henton & Bladon 1988, Ogden 2001), there have yet to be any studies of creak in the Northwest. This paper is an investigation into Oregonians' use of creak, utilizing interviews with four speakers. The interviews were transcribed in ELAN, and evaluated using Praat and statistical analysis of individual and collective rates of usage. The results show that creaky voice tends to occur in clause-final position, and in extended turns, suggesting differences in the phonology of Oregon English, or in the function of creak, such that it signals pauses in speech.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF EXAMPLES

LIST OF FIGURES

1. Introduction

1.1 The Pacific Northwest

A common belief in the United States is that natives of the Pacific Northwestern region of the United States (typically defined as Oregon, Washington, and Idaho, occasionally extended to include British Columbia in Canada) "don't have an accent". Although most US Americans would agree that Californians speak with some sort of discernible accent (due in no small part to caricatures of surfers, movie stars, and other stereotypically Californian characters often portrayed in film, television, and literature) it is widely believed that the English spoken in the Northwest is neutral. This view of Pacific Northwestern English is by no means limited to out-group perceptions of the variety, rather most Oregonians or Washingtonians will themselves insist that "of course we don't have an accent!"[1] This is, of course, absolutely contrary to what is taught in even the most basic course in linguistics, where it has long been understood that *everyone* has an "accent". This discrepancy between the fundamental tenets of linguistics and popular views of language has everything to do with the way "accent" is defined. To the layman, an "accent" is something that departs from the standard in one or more ways (see Lippi-Green 1997 for a comprehensive discussion of this), whereas in linguistics, if the term is used at all, an "accent" is simply a shorthand way of referring to the phonetics and phonology (and sometimes also the syntax, semantics, etc) of a given variety, independent of any standard. Unfortunately, although we can be sure that

[1] In fact, according to Hartley (1999), Oregonians rate their own English as the most "correct" English in the US.

Northwesterners have an “accent” just as speakers from every other region do, the exact nature of the dialect spoken in the Pacific Northwest remains unclear.

Pacific Northwestern English, as part of the “third” dialect of US English, has seen relatively little scholarly investigation in contrast with other, more established dialect areas, even when compared with other regions in the West, such as California. Most often, the West is defined in relation to other major dialect areas in North America: the absence of “Canadian” Raising in /aw/ distinguishes it from English in Canada, the low back merger from the North East, absence of /ow/ fronting from the Midlands, and the absence of glide deletion in /ay/ from the South (Labov, Ash, & Boberg 2006). Despite lexical (Vaux & Golder 2002) and phonological (Conn 2002, Ingle, Wright, & Wassink 2005, Ward 2003) support, there have been few efforts made to carve up the West into smaller dialect areas, even though some scholars have suggested that the Pacific Northwest, as a discrete dialect area, is the most distinct in the West (Ward 2003 p. 35). This lack of research stems in part from the fairly young age of the dialect area, and indeed, the state of Oregon, the focus of this study, is only just over 150 years old, with Washington State being younger still. Large-scale migration to Oregon began around 1840, with immigrants coming from all around the country, but primarily from the American midlands (Pennsylvania, Ohio, Indiana, Illinois, Missouri, Kansas, Nebraska, and Iowa), and in 1859, Oregon was officially inducted into the union. Presumably the array of different varieties spoken by immigrants coming to Oregon resulted in dialect contact and subsequent leveling, with the resulting mixture becoming

the progenitor of the English now spoken in the region[2] (Trudgill 2008). The age of the region means that English in Oregon has only had around 150 years to develop divergent features or to participate in large-scale chain-shifts and other changes going on around the continent. Oregon's young age presents a special problem to researchers, namely that of finding "native" Oregonians. According to the 2000 US Census, only 45.3% of Oregon's population was born in Oregon, with the rest having immigrated mainly from elsewhere in the US or Canada (US Census 2005). A consequence of this is that it is very difficult to find people who can trace their family back more than two generations in the Pacific Northwest: most are 1st or 2nd generation. In fact, in the 2000 census, only 544 people listed their ancestry as Oregonian (although this may not be entirely representative due to non-responses, people self-reporting simply as US American, and people identifying with their parents' place of origin, their ethnicity, or their ancestors' place of origin) (US Census 2007). The scarcity of true "natives" means that if these 2nd and 3rd+ generation northwesterners do have any unique features, they may not be spreading simply because there are too few native speakers, and too much input from immigrants from other dialect areas (to say nothing of the leveling effects of modern communication and transportation). The diverse array of people who have immigrated to Oregon means that there are a large number of dialect features which could act as potential input to people growing up in the area. Given the constituency of the state then, the challenge should perhaps be couched in terms of identifying emergent features characteristic of

[2] Because immigration patterns were quite similar throughout the West, it should not come as a surprise that the different dialects of English in the Western US share a number of features.

Oregonian English in the broader sense, rather than uncovering the attributes of "true" Oregonians, who would appear to be in the minority.

Despite the challenges facing researchers in the area, there have been some studies of Pacific Northwestern English – chiefly in Washington State and in Portland, Oregon – whose results suggest that some distinctive characteristics are emerging. These studies have noted differences in the articulation of vowels, such as the tendency for front and back vowels to be centralized (Conn 2002, Ingle et al. 2005, Ward 2003), along with the presence of some features known from other dialect areas, such as positive *anymore* (J. Conn, personal communication, February 9, 2009). Although these features of Pacific Northwestern English are not common knowledge outside the realm of linguistics, there are some well-known lexical differences (Oregonians don't go to the *sea* or the *beach* or even *down the shore*, they go to the *coast*; etc), some due to contact with indigenous languages, especially for example, via Chinook Jargon (Conn 2006).

1.2 Creaky Voice in the Pacific Northwest

One feature of Pacific Northwestern English that has often been reported in studies of the variety is the prominent use of creaky voice (Ingle et al. 2005, Ward 2003). Creaky voice is a type of phonation characterized by a low pitch and a slow rate of vibration, reminiscent of the sound made by a creaky door; whence the name. Although much of the recent literature on Pacific Northwestern English mentions creaky voice (cf. Ingle et al. 2005, Ward 2003), there are as yet no studies, quantitative or qualitative, of its function within the dialect area. Although creaky occurs in nearly everyone's speech at

certain times, due to some physiological factors which condition its occurrence, if it is indeed more common in the speech of Pacific Northwesterners, then that would suggest that there is more than physiology at work. Looking cross-linguistically and even cross-dialectally, different phonation types such as creaky voice (Henton & Bladon 1988, Ogden 2001, Podesva 2007) and falsetto (Podesva 2007), along with other features such as ingressive airstream (Clarke & Melchers 2005) have been shown to be relevant at the level of discourse. The fact that there is a precedent for phonation serving a discourse function in other languages and dialects of English means there is a very real possibility that the use of creaky voice is linguistically significant in this dialect. This paper is an investigation into the function of creaky voice in Oregon English, utilizing two recorded interviews with four Oregonians (made for another paper on Oregon English). The interviews were transcribed using ELAN annotation software, after which all tokens of creaky voice were extracted, coded, and evaluated using basic statistical analysis and Praat phonetic analysis software.

2. Literature Review

2.1 English in the Pacific Northwest

Aside from two articles written in the 1960's (Foster & Hoffman 1966, Reed 1961), and one in 1980's (Reed 1983) – all of which are largely obsolete today – there has been very little work done on Pacific Northwestern English. Foster and Hoffman (1966) is an overview of Pacific Northwestern English which looks at the vowels of speakers in the Seattle area of Washington State. Foster and Hoffman (1966) comment on the low-back merger, as well as the fact that words such as *tune*, *new*, and *due* tend to be pronounced with [u] rather than [ju] in PNW English (though this is common throughout the west, as well as other dialect areas).

Another early study was Reed's 1961 article "The Pronunciation of English in the Pacific Northwest", which is also a brief survey of Pacific Northwestern English, including a listing of all the vowels and consonants found in the variety. Reed (1961) also observes that local speakers appear to have the low-back (*caught*/*cot*) merger, as well as post-vocalic /r/. Additionally, he notes some vowels and consonants which are in variation in the region, however many of these – such as /hw/ – are no longer relevant to contemporary studies (Reed 1961).

It is only in recent years that scholarship focusing on English in the Northwest has begun to pick up, beginning with the Portland Dialect Study, which was completed in 2003 at Portland State University. The Portland Dialect Study includes over 60 informants, of various ages, sexes, and social classes, who were interviewed for approximately 30-45 minutes (Portland Dialect Study 2003). The results were subjected

to instrumental analysis, and made available to students of linguistics at PSU for research purposes. A number of theses came out of the Portland Dialect Study, including some on topics such as high rising contours Portland English, and /æ/ in Portland English. One thesis, "Portland Dialect Study: The Fronting of /ow, u, uw/ in Portland, Oregon", written in 2003 by Michael Ward, is a comprehensive study of back vowel fronting in the speech of Portland natives. Ward's (2003) data show strong, statistically significant degrees of fronting, correlated most closely with age (using the apparent time construct) with the youngest speakers showing the highest rates, correlated rather weakly with gender and social class, with women and the working class leading slightly. He concludes by suggesting that more work be done on what sorts of dialect borders exist in the west, as divisions within the "third dialect" are currently unclear (Ward 2003).

More recent studies have been undertaken by Alicia Beckford Wassink at the University of Washington, and Jeff Conn at Portland State University, who jointly run the Pacific Northwestern English Project, a large-scale study of speakers from the Pacific Northwest. Although results from the large-scale study being undertaken by Conn and Wassink are not yet available, both Conn and Wassink have written, contributed to, and presented, papers on English in Portland and the Pacific Northwest in general.

Conn's 2002 paper, presented at NWAV 31, is one of the best sources of information on contemporary Pacific Northwestern English. Conn (2002) uses interviews with 14 informants (subsuming three age groups, both sexes, and two social classes) from the Portland Dialect Study and looks at the degree of back vowel fronting (specifically /u/, /uw/, and /ow/) in the speech of each informant, as well as the low-back merger, and

the distribution of /æ/. Conn's results show back vowels with fully front realizations, although not without some variation. With the exception of the tokens which are nasal-adjacent or which precede a lateral, nearly all instances of /u/ have an F_2 of 1400 Hz or above, but the F_2 values for /o/ range from around 850 to 1500 Hz (Conn 2002). The low-back merger appears to be complete for the younger speakers in Conn's (2002) data, however one of his older informants, an 89 year old woman, did have two distinct realizations of the vowels. Conn compares his results to data on the Canadian and Californian shifts, and concludes that his informants are indeed participating, but that traditional indices of speaker class do not seem to be linguistically relevant in Portland – that is, they do not capture any patterns in the data. Perhaps, he says, this is due to inflated housing costs and the growing percentage of people under 40 years of age with post-secondary education, but more research on this topic and a larger pool of informants are needed.

Another recent study is Ingle, Wright, and Wassink's "Pacific Northwest Vowels: A Seattle Neighborhood Dialect Study" (2005), which is a phonetic analysis of the vowels and speech of informants from the neighborhood of Ballard in Seattle. The informants in the study were recorded reading a word list targeting the 15 General American vowels – which tends to elicit more formal pronunciations – and while in conversation with another person matched in age, social class, etc – which tends to elicit more informal speech. Ingle et al. (2005) found that the women they interviewed were much more likely to use creaky voice (6 out of 8) than the men in their study were. The data also show that the informants in the study appear not to be fronting /u/ as much as

has been reported in California (Ingle et al. 2005). Interestingly, Ingle et al. (2005) also found that some speakers appeared to be participating in the Northern Cities Chain Shift – a shift which, according to other researchers, is not occurring in the Pacific Northwest (Labov et al. 2006) – and that the speakers did not appear to be participating in Canadian Raising.

A more general source which is of use when investigating North American dialects of English is Labov, Ash, and Boberg's *Atlas of North American English*. The *Atlas* has relatively little to say about the Northwest, or the West in general, stating chiefly that it is a developing dialect area without sharply defined boundaries, defined by a fronted /uw/, a non-fronted /ow/, lack of participation in Canadian Raising, and lack of glide deletion in /ay/ (common in the US South) (Labov et al. 2006 p. 279). The *Atlas* does have eleven recorded speakers in the Pacific Northwest (specifically in Washington, Oregon, and Idaho), but only four of these speakers are actually from Oregon. Data from the *Atlas* show a fronted /u/ (F_2 values between 1600 and 1800 Hz), and a moderately fronted /o/ (F_2 between 1200 and 1300 Hz) (Labov et al. 2006). For comparison, the Atlas lists Midland values for /o/ as being far more front (around 1400 Hz, with some even greater than 1550 Hz) (Labov et al. 2006).

The Dialect Survey, a website run by Vaux & Golder, has more information from the Northwest, although their methodology differs substantially from that used by Labov et al. (2006). In collecting the data for their website, Vaux & Golder created a 122-question survey, soliciting responses from people around the US, the results of which are displayed on the website. A key difference in methodology between this project and the

Atlas of North American English is the fact that Vaux & Golder used a written survey, explicitly asking informants about their speech (Vaux & Golder 2002), whereas Labov et al. (2006) used telephone interviews with respondents. Although there are many problems with asking informants to self-report, surveys have the advantage of being able to cover a much wider area with much less effort, and as such Vaux & Golder (2002) had 166 responses from residents of Oregon, compared to the aforementioned 4 (Labov et al. 2006). Data from the Dialect Survey largely concern lexical choice, although some deal with vowel choice in certain words. Unfortunately this data is of little for the present study, as the respondents were not asked about their perceptions or use of creaky voice.

My own study of English in the Pacific Northwest (which the interviews analyzed in this paper were recorded for, see section 3) looked at the vowels of Oregon English (Riebold 2009). Five tokens of nine simple vowels ([i, ɪ, e, ɛ, æ, a, o, ʊ, u]) from each speaker were extracted from the interviews with Sony Sound Forge, and were analyzed using Praat in order to measure formant values. The tokens were selected based on the context they occurred in, and to that end, all tokens which were adjacent to a nasal were excluded, as well as those preceding a lateral, as both of those environments are known to inhibit fronting (Conn 2002). The five tokens of each vowel from each speaker were averaged, and plotted by using formants 1 and 2 as indicators of the height and frontness/backness of the vowels. The results show a tendency towards centralization in the speech of all four correspondents, with strong fronting of /u/ (F_2 > 1600 Hz) and moderate fronting of /o/ (F_2 > 1100 Hz), as well as backing of front vowels (e.g. F_2 for [i] < 2100 Hz) (Riebold 2009).

2.2 Creaky Voice

The type of phonation referred to here as creaky voice goes by many names: glottal fry, vocal fry, pulse phonation, laryngealization, and creaky phonation. Not everyone considers these terms synonymous however, and some linguists (see Laver 1994) make a distinction between creaky *voice* and creaky *phonation* (or simply creak), where the former is a compound phonation composed of creak and voicing, and the latter is a voiceless creak. Following Laver (1994), I will not make use of this distinction, as it has never been shown to be linguistically significant, and it is doubtful that the two components of creaky voice are indeed separate and used as such; therefore I will treat the various terms as synonymous for the purposes of this paper. Although this paper does not concern itself with the precise articulatory and acoustic details of creaky voice, a brief discussion of these is beneficial to the understanding of the phenomenon as a whole.

Auditorily, creaky voice is a phonation characterized by a slow "creaking" sound (Gordon & Ladefoged 2001), wherein it almost seems possible to perceive the individual glottal pulses which make up the phonation. The noted phonetician Ian Catford famously compared creaky voice to the "a stick being run along railings" (Catford 1964 in Laver 1994).

2.2.1 Articulatorily

Although nearly everyone agrees on what creaky voice sounds like, exactly how it is articulated is disputed (cf. Henton & Bladon 1988, Laver 1994). One theory is that

during the production of creaky voice the vocal folds are relaxed slightly and bunched together, resulting in a “thick vertical contact” which damps vibration of the glottis, and leads to what we perceive as creaky voice (Laver 1994). Others suggest that the low frequency pulses are produced when only a small length of the folds at the front of the glottis vibrates (Catford 1964 in Laver 1994, Ladefoged 2001), and in fact this assertion is supported by photographs of the glottis in action. Other features such as subglottal air pressure are contested, with some maintaining that subglottal pressure is very low (c.f. Henton & Bladon 1988, Laver 1994), and others suggesting that subglottal pressure may be at its highest during the production of creak or falsetto phonation (cf. Henton & Bladon 1988). Despite this disagreement, it seems clear that subglottal air pressure isn’t critical to the production of creaky voice, and as such it is merely a secondary concern. Laryngoscopic photographs of the glottis also show alternation between single and double glottal pulses, between which vibration is completely damped, however it is also doubtful that this pulse grouping has any effect on the perception of creaky voice. Regardless of how creaky voice is articulated, it leads to a reduction in glottal vibration, and a significantly reduced volume of air flowing through the vocal tract (Henton & Bladon 1988, Laver 1994).

2.2.2 Acoustically

Acoustically, creaky voice features an F_0 frequency of between 20-70 Hz, which is well below typical modal voice frequencies (Henton & Bladon 1988, Laver 1994). The glottal pulses are aperiodic and separately resolvable both auditorily and under

instrumental analysis. On spectrograms, the pulses appear as dark, vertical striations, at lower overall amplitudes, and the irregularity of those pulses is belied by the spacing of the striations. Similar to the debate surrounding what characterizes creaky voice in articulatory terms, the acoustic classification of creaky voice is also subject to disagreements and protracted debates. Although many people identify the low F_0 frequency typically found in creaky voice as being a defining feature of it (Laver 1994), there is ample evidence showing that F_0 alone cannot account for a stimulus being perceived as creaky (Gobl, Bennett, & Ní Chasaide 2002). Gobl, Bennett, & Ní Chasaide (2002) found that the manipulation of F_0 alone was not enough to adequately signal a given voice quality, though they note that this is perhaps due to the unnaturalness of stimuli which have had their F_0 modified, but are otherwise identical.

2.2.3 Causes

The occurrence of (non-contrastive) creaky voice is often physiologically conditioned, and occurs spontaneously in almost all speakers, as a relaxation of the vocal folds can be the result of physical exhaustion or drowsiness. Relaxants (such as alcohol and some narcotics) can also trigger a slackening of the vocal folds, and even being sick is sometimes enough to cause creaky voice (Laver 1994). Glottal segments (e.g. glottal stop in English vowel-initial words) also tend to co-occur with creaky voice, and vocal fold elasticity (the tendency for vocal cords to “snap” back after having been stretched) can trigger creaky voice if a speaker reaches falsetto or near-falsetto pitches (Podesva 2007). Finally, and perhaps most important to this study, creak phonation can be

prosodically triggered, for example by English intonational units which tend to end with a fall in pitch. As creaky voice is almost always produced with a very low F_0, a steep enough drop in pitch is capable of triggering creaky voice (Henton & Bladon 1988), although it is likely that there are other factors involved.

2.2.4 Usages

As with most linguistic features, creaky voice is used in a variety of ways cross-linguistically, contrastively, as well as discursively. For example, it has long been known that creaky voice is used contrastively in some languages, such in Burmese, Vietnamese, and those in the Nilo-Saharan language group (Henton & Bladon 1988), as well as Danish and some Chadic languages (Laver 1994). In languages such as these, creaky voice contrasts with modal voice, and occasionally also with breathy voice. Interestingly, contrastive creaky voicing has been reported in some Pacific Northwestern aboriginal languages, such as Spokane (Carlson & Esling 2000), and creaky voice has been found to be conditioned by glottal segments in Nuuchahnulth (Carlson, Esling, & Fraser 2001), although it is very unlikely that there is any connection between the use of creaky voice in these languages and Pacific Northwestern English. As mentioned above, contact between English-speaking settlers and indigenous tribes was frequent, as evidenced by the existence of Chinook Jargon, but this is also presumed not to have been a factor in the development of any local usage of creaky voice in Pacific Northwestern English, as it is clear that creaky voice is not lexically contrastive in this variety.

Podesva's (2007) paper focuses on phonation as a stylistic variable. Podesva (2007) draws his data from the speech habits of one man, Heath, who, according to him "exhibit[s] considerable cross-situational variation". Podesva contrasts Heath's use of falsetto with his use of creaky voice, using data recorded in the number of different social situations. After a short discussion of the factors which typically condition the occurrence of creaky voice, Podesva concludes that these factors alone cannot account for all of the tokens, and notes previous studies which have found a connection between the use of creaky voice and social factors (see discussion of Henton & Bladon 1988 and Pitam 1987 below). The analysis of Heath's usage of creaky voice presented is that the two types of phonation (creaky and falsetto) are used in tandem to widen Heath's F_0 range, noting that these phonations are used with greater or lesser frequency depending on the situation. Podesva argues that Heath's identity as a gay man conditions his use of falsetto, which he uses to create a "diva" persona, whereas his use of creaky voice is a way of throwing that falsetto into sharp relief. In this way, the subject's usage of different phonations contributes to the creation of a dynamic identity, which changes from situation to situation.

Moving away from English, other studies of creaky voice have focused on its role in turn-taking systems. In Finnish, for example, Ogden (2001) argues that creaky voice is used to signal the end of a turn-at-talk. Ogden uses a corpus of data drawn from a program on Finnish national radio wherein listeners call in to request particular songs, and to discuss the songs themselves. The corpus consists of the transcriptions of nine calls, which Ogden analyzes within the framework of conversation analysis. Ogden notes

that creak is often found at the end of complete Turn Constructional Units (TCUs), in parenthetical utterances, word-medially, over short stretches of speech, and at Transition Relevance Places (TRPs). When creak occurs turn-finally, it most often occurs following voiceless obstruents and on syllables which do not bear primary stress (Ogden 2001). Ogden (2001) finds that, overall, creaky voice is most often employed to signal the end of a turn, showing several extracts from the data exemplifying this. The data also appear to show glottal stops being used to extend turns, an assertion for which Ogden also gives supporting examples. Ogden concludes by suggesting that even if different phonations are not lexically contrastive, they can still play an important role in interaction.

Additional studies have shown that Finnish speakers of English carry over this feature of their native language into their L2. Toivanen & Waaramaa (2001) found that Finns expressing mitigated disagreement in English tended to end their TCUs with creak, although in contrast to Ogden (2001), they note that creaky voice appears to be related to attitude more than to turn-taking.

2.2.5 Social Evaluations of Creaky Voice

The use of creaky voice is quite salient, and as an idiolectal feature it is subject to a number of value judgments similar to those surrounding high rising terminals in "Valley Girl" speech (as well as the speech of many other age groups and dialect areas). Like high-rising terminals, opinions of creaky voice are often negative, characterizing it as a trait to be gotten rid of, rather than a facet of one's idiolect or dialect. Also similar to high-rising terminal contours, many people (cf. Laver 1994) characterize it as

representative of some aspect of the speaker's personality, for example, a lack of confidence or some other personality "flaw". Beyond simply having a low opinion of it, some actually go so far as to classify it as a dysphonia, or a speech pathology (Henton & Bladon 1988), and even within the field of linguistics there are negative social evaluations to be found, such as John Laver's assertion that creaky voice was emblematic of "bored resignation" (Laver 1994 p. 196).

In reaction to this, some studies have attempted to quantify speakers' judgments of creaky voice. Pittam (1987) details an experiment measuring Australians' reactions to several different types of phonation: breathy, creaky, nasal, tense, and whispery voices. Pittam's (1987) results show that while male tense voice tends to be rated highly on the status scale, and female breathy voice high on the solidarity scale, the rankings for creaky voice were mixed. Creak voice ranked lowest on the "solidarity" scale, and in the middle on the "status" scale, finding that there were no major age or sex effects, although the results do also suggest that creaky voice is associated more with men than women, perhaps due to the very low F_0 frequencies that characterize it (Pittam 1987).

A more recent study was conducted by Gobl and Ní Chasaide (2003), entitled "The role of voice quality in communicating emotion, mood and attitude". Gobl & Ní Chasaide documents listeners' reactions to samples of synthesized speech – based on a recorded Swedish utterance – designed to emulate seven different voice qualities (harsh, tense, modal, whispery, breathy, creaky, and lax-creaky). The utterances were generated using a formant synthesizer, and the exact settings were based on results from prior studies and impressionistic judgements. The results of the study show creaky voice and

lax-creaky voice (a combination of creaky voice and breathy voice) being correlated with "low-activation" emotions, such as, interestingly enough: boredom. This correlation is far from a one-to-one mapping though, as creaky voice is also rated highly for relaxed, intimate, content, sad, and friendly, and is therefore not limited solely to negative judgements. Lax-creaky voice tended to be more strongly associated with the aforementioned traits, although it is unclear if this mode of phonation is in fact distinct from garden-variety creak. In their conclusion, Gobl & Ní Chasaide note that although the synthesized voice qualities used in their experiment seemed effective at signalling mildly affected states, it is doubtful that these stimuli alone are being use to signal any one of the moods identified. (Gobl & Ní Chasaide 2003)

2.2.6 Creaky Voice in English

Henton & Bladon (1988) is a study of the use of creaky voice in a corpus of data collected from speakers of two dialects of British English: Received Pronunciation (RP) and Modified Northern[3] (MN). Henton & Bladon set out a number of research questions to be investigated in their paper: how much creaky voice is used in a large corpus of natural speech, is creak more common in utterance-final syllables with falling-intonation, does the use of creaky voice vary according to the sex of the speaker, does it vary according to dialect, and do non-pathological speakers use creaky voice exclusively. Another aim of their study is to address a number of shortcomings which the authors identify as characteristic of previous studies, namely: impressionistic data, small numbers

[3] Defined by Henton & Bladon (1988) as speakers from northern Britain who, in the course of their lives, have relocated to another part of the country and presumably "modified" their speech.

of speakers who are deliberately asked to use creaky voice, failure to consider regional and sociolinguistic variation, exclusion of female speakers, and "artificial" speech. The corpus under analysis is the Oxford database, which was recorded as part of a project investigating British English vowels. The database contains 40 speakers: 10 women and 10 men from each dialect. The speakers were asked to record contextualized sentences containing /hVd/ words in a random order, and were also directed to place stress on the names included in the sentence, so as to control for the effects of stress on vowel production. Henton & Bladon note three key advantages to using this corpus: the sentences are "naturalistic" and varied, they were composed for a different purpose, and they were the first exercise in the series (this is meant to address Laver's claim that creak indicated "bored resignation"). Henton & Bladon listened to the 880 sentences contained in the corpus three times individually, then once more together. All syllables in the database (excepting those from one speaker, whose recording was degraded and was therefore excluded from the analysis) were ranked based on the amount of creak they featured (scale not given). The tokens were also divided into four classes, final syllable (FS), pre-final syllable and final syllable (PFS + FS), other syllable (OS), and all syllables. Additionally, the authors created spectrograms of various tokens as a means of verifying their judgments, and so as to be able to compare creaky voiced syllables with modal voiced syllables. The results of the study show that final syllable position has a strong, statistically significant ($p<0.01$) effect on whether speakers use creaky voice or not, accounting for 207 out of 548 tokens (37.7%). Following close behind, PFS + FS accounts for 167 tokens (29.5%). Other results of the study are that MN speakers use

creaky more than RP speakers, and that men use creak more than women. This last finding contrasts with reports of creaky voice in Pacific Northwestern English, which commonly note creaky voice being more common among women (cf. Ingle et al. 2005). Henton & Bladon also discuss the possibility of creak being a marker of androgyny (for men to sound less masculine), or as a means of differentiating men from women (making men sound more masculine). The last point addressed by Henton & Bladon is whether non-pathological speakers habitually use creak throughout their utterances, and in fact the authors found that 10 out of the 79 used creaky voice more than 65% of the time. Although this is an extremely useful paper, particularly as a foundational approach to studies of creaky voice, there are a few aspects of it which are in need of improvement. First, the analysis of what the authors call "Modified Northern" is puzzling, as it does not, in all likelihood, represent a true dialect, rather some sort of "interdialect" variety between Northern British English and RP. From the results, one may infer that Northern British speakers use creak more often, however that is hardly conclusive from the data, given that these speakers are supposedly not representative of true Northern British English. Secondly, although the corpus includes a large number of speakers, it is "naturalistic" rather than natural data, in so far as it is elicited rather than spontaneous. Finally, in restricting tokens to one or two syllables and making a distinction between PFS + FS and FS, Henton & Bladon are perhaps missing a generalization. It is not clear that FS and PFS + FS should be different environments, and indeed when combined, the superordinate category of utterance-final accounts for 369 out of 548 tokens, which is 67% of the tokens.

Laver (1994) also mentions the use of creak voice in English, stating that creaky terminations signal TRPs for some speakers. Unfortunately the source or supporting data for this assertion is unclear (no reference is given), although many have since cited this claim.

A final mention of creaky voice comes from a study of Utah English in 1990 (Di Paolo & Faber 1990) where it was found that front vowels tended to be pronounced with creaky voice, with the probability of creak increasing for lower vowels. From the data, it appears that speakers of Utah English may be using creaky voice in order to distinguish lax vowels from tense vowels, with lax vowels favoring creak. Although Utah is not generally considered part of the Pacific Northwest, it is part of the third dialect of US English, and therefore has some relevance to this study, even if only for the purposes of establishing a precedent for creaky voice being linguistically relevant in dialects of English.

2.2.7 Creaky Voice in the Pacific Northwest

As I have already alluded to, creaky voice in Pacific Northwestern English has been mentioned by several researchers in the area, but has never studied in its own right. Thus far the occurrence of creaky voice in this dialect of English has been mentioned in one of the key papers on Pacific Northwestern English (Ingle et al. 2005) and by Jeff Conn, another researcher in the region (personal communication, February 9, 2009). Ingle et al (2005) note that although it is not unusual for US American males to use creaky voice, it is typically less common among women. Ingle et al. go on to say that

despite this, women in the Pacific Northwest appear to use creaky voice more than those in other dialect areas, and that creaky voice showed up even in the elicitation tasks they carried out as part of the research project. Ingle et al. also provide two example audio clips taken from the interviews recorded for the paper, both of which feature marked use of creaky voice, as strong or stronger than that used by my informants.

3. Methodology

3.1 Interviews

The two interviews analyzed in this paper were carried out in groups of two, so as to facilitate dialogue between the informants (rather than monologues directed towards the interviewer) and thereby reduce, if possible, the effect of the observer's paradox. Each interview was approximately one hour long, with one or two prompts given initially in order to start the conversation, after which a more or less natural flow of topic shifting commenced. The data analyzed were taken from natural speech, with no word lists or elicitations used.

The interviews were recorded to uncompressed, 44 KHz, 16 bit WAV files using a Shure SM58 vocal microphone plugged into an E-MU 0202 audio interface, which was connected via USB 2.0 to a laptop computer. The interviews were held in familiar, informal, and quiet locations, in order to reduce any effect the setting (such as a soundproof booth) could have on the speech of the informants, and in order to prevent any outside noise from distracting the participants. The only people present during the interviews were the two speakers and myself, the interviewer.

3.2 Participants

Of the four speakers in this study (three males and one female) three were born in Corvallis, a small town in the Willamette Valley about two hours south of Portland.

Speaker 4 moved to Corvallis from California at approximately 5 years of age[4]. All but one (speaker 4) have been living elsewhere in Oregon since they were approximately 18 years old, although this is unlikely to be a factor, as there is thus far no evidence that English in Oregon differs markedly from one area to the next. The breakdown of the speakers and interviews are illustrated in the table below:

Interview	Speaker	Age	Sex
I	1	24	Male
	2	22	Female
II	3	23	Male
	4	24	Male

Table 1: Speakers

As should be clear from the table, the informants are all relatively close in age, and are also all members of the same social class (middle) and race (white). Informants 1, 2, and 4, are university students, and informant 3 has already completed a university degree, and is on his career path in journalism. Such a uniform selection of informants is intentional, because it maximizes the significance of the results obtained. If the informants were stratified, say on the basis of class, with each being from a different social class, it would be impossible to judge whether any of them were representative of their class or not. Therefore, because the study has only four informants, having them all be from a single region, social class, and level of education, maximizes the amount of relevance the data has.

[4] 5 years old is typically believed to be the age after which a person cannot acquire a new phonology (Labov 2006), and in the paper I wrote on Oregon English, this speaker's vowels did in fact match those of the other speakers. I therefore include him in this study as well, barring any evidence that his dialect differs significantly from that of the other speakers.

All four speakers have studied at the university level, meaning it is possible that their speech may be more standard than that of other speakers, such as those with less post-secondary education or from a lower social class. The impact of this post-secondary education may be lessened due to the widespread belief that Oregon English is very close to the standard, that is, because there is little to no pressure put on students to acquire a more standard variety (apart from some prescriptivist norms) the speakers are less likely to have consciously modified their speech.

3.3 Analysis

The interviews were transcribed using ELAN, with a separate tier used to mark each speaker's use of creaky voice within a given utterance. What does and does not constitute an utterance is partially contingent on the way ELAN handles annotations. Because ELAN works by allowing the user to select sections of the waveform which correspond to stretches of speech, and which can then be annotated, the transcriber is often presented with the choice of whether to have two separate utterances for a given section of speech, or one. Depending on how the data is used, this may be a trivial concern, as ELAN allows the user to generate any number of statistics regarding the transcript, such as word count, percentage of time spoken, etc. In this paper however, because utterance position has been identified as a potentially relevant factor, the definition of "utterance" is more important. For my purposes, I attempted to separate stretches of speech based on pauses and prosody. If there was a pause of an appreciable length (>500 ms), or if the speaker paused to initiate word-search, I would begin a new

utterance. If the speaker was mid-sentence and the silence represented an in-breath, and if the prosody suggested an incomplete utterance, I transcribed the speech as one larger utterance.

Creaky voice was first identified impressionistically, and the resulting tokens were spot-checked in Praat following the criteria set down in the literature (cf. section 2.2, section 4, Gobl & Ní Chasaide 1992, Henton & Bladon 1988, Laver 1994, Moosmüller 2001). A token is defined as an unbroken segment of creaky voice, and may therefore range in length from one syllable all the way to an entire utterance. This definition of the token diverges from that adopted by Henton & Bladon (1988), who limit their tokens to one syllable in length, with the exception of PFS + FS, which is two syllables in length. The motivation for this difference in methodology comes from a basic observation about my data, namely that when creaky voice is used, it is seldom confined to a single syllable, rather it typically ranges over a single (polysyllabic) word, or several words (see section 4.2 for details). This difference in the length of creaky voice tokens may also speak to differences between British English dialects and North American, or Pacific Northwestern dialects of English, however it may also be a purely methodological difference. Each annotation on the phonation tier records the entire utterance, with the portions featuring creaky voice demarcated by tildes, similar to Laver's proposed transcription convention (1994 p. 196-197). After transcription, all tokens were extracted, and the number of words spoken with creaky voice was counted for each speaker. Additionally, the tokens were coded for their clausal position (initial, medial, final, and entire), utterance position, and turn position, as well as for the presence of an adjacency

pair, a discourse marker, or a filler. It should also be noted that there were a few examples of creaky voice which were not counted as tokens, namely those where creaky voice is used in reported speech. Examples of this are rare, only occurring a handful of times in my interviews, however they are noticeably different from other uses of creaky voice, in that they only occur when one speaker is imitating another, and strong creaky voice is used consistently throughout the imitation.

Speaker 4: "we'd take turns like DJing basically or whatever and we'd just play stuff and you know I always like '~smooth it over~' and played some like R&B crap and like..."
[…]
Speaker 3: "we would just like, yeah you would just go and be like..."
Speaker 3: "'alright this next track, ah, I'm just going to ~smooth it over a little bit~' and it was like, after every song you'd say the same thing"
Speaker 4: "yeah because, uh, I mean I did it twice and it became a joke"

Example 1: Creaky Voice in Reported Speech (II / 58:31.710 - 59:38.230)

This example is clearly different from other instances of creak, as speaker 4 uses significant creaky voice when repeating a phrase he used to use when DJing an internet radio station, and speaker 3 repeats the phrase a short time later with an equally heavy amount of creak while imitating speaker 4. Since, in this example, the creaky voice seems to be an integral part of the performance, as well as part of an inside joke, it is presumably not optional, and therefore is not expected to pattern along with other tokens of creak, or indeed pattern at all. Instances such as these (and there were very few) were not included with the other tokens, as they don't represent the usages being investigated in this study.

Once extracted, the tokens were analyzed using basic methods of statistical analysis, such as the calculation of frequencies, averages, and percentages. I feel that the data gathered in this project, as well as the size and design of the project itself, do not lend themselves to the multivariate statistical analysis techniques common in much of the research on topics relating to sociolinguistics for several reasons. First, using multivariate statistical analysis, it would be necessary to tokenize and code all non-creaky utterances in the data as well. However, due to the wide range of environments in which creak can occur (fillers, clause boundaries, adjacency pairs, etc), it would be impossible to define a consistent environment for tokens in which creak is absent. Similar problems have been encountered in other sociolinguistic studies of suprasegmental features, which have also refrained from using multivariate analysis (Podesva 2007). Furthermore, in coding the tokens for analysis, it would be difficult to isolate factors that directly relate to physiology, rather than being simply linked to physiology by way of English prosody (such as the utterance-final environment). Because multivariate analysis offers relatively few advantages for this project, and because more basic methods of statistical analysis are more efficient for my purposes, I chose to make use of the latter.

3.4 Limitations

Although all efforts have been made to be thorough in the analysis of the data, the fact that the data is from only four speakers must still be taken into account. The conclusions drawn from this study are therefore necessarily preliminary, and may only be relevant for a small subset of speakers in Oregon or the Pacific Northwest. Additionally,

as has already been mentioned, due to the relatively small scale of this project, the participants are representative of only one class and race. It has been claimed that, in the Pacific Northwest, women use more creaky voice than men, and indeed this does appear to be the case based solely on my impressionistic observations made while traveling around Oregon. Unfortunately I cannot substantiate this claim with my data as only one of the speakers is a female, meaning that barring any further research on this topic, it would be difficult to prove that her usage of creaky voice is indeed representative of women of her age, social class, etc.

Regarding the analysis in particular, it is often the case that two or more people coding the same tokens will not agree on every categorization. Many experiments deliberately confront this issue by asking multiple researchers to code and discuss the tokens, so that an agreement can be reached. With this paper, however, I am the sole coder, and therefore it is possible that I have made mistakes in this endeavor, although I have taken great pains to check over all the data multiple times.

Lastly, the definition of "utterance" presented above is problematic. First, it hinges upon an arbitrary amount of time (500 ms) in order to separate utterances, but is made necessary by the workings of the transcription program itself. The separation of the speech signal into discreet utterances has significant ramifications for this data, as separating utterances could divide an otherwise continuous token of creaky voice. Although utterances are clearly important in other contexts, they are perhaps too artificial for the purposes of this paper.

4. Results

Echoing the descriptions found in Laver (1994) and Henton & Bladon (1988), creaky voice as used by the speakers in these interviews is clearly visible in spectrograms. The following spectrograms (created in Praat) illustrate this difference – visible even within ELAN, which offers a considerably zoomed-out view of the waveform.

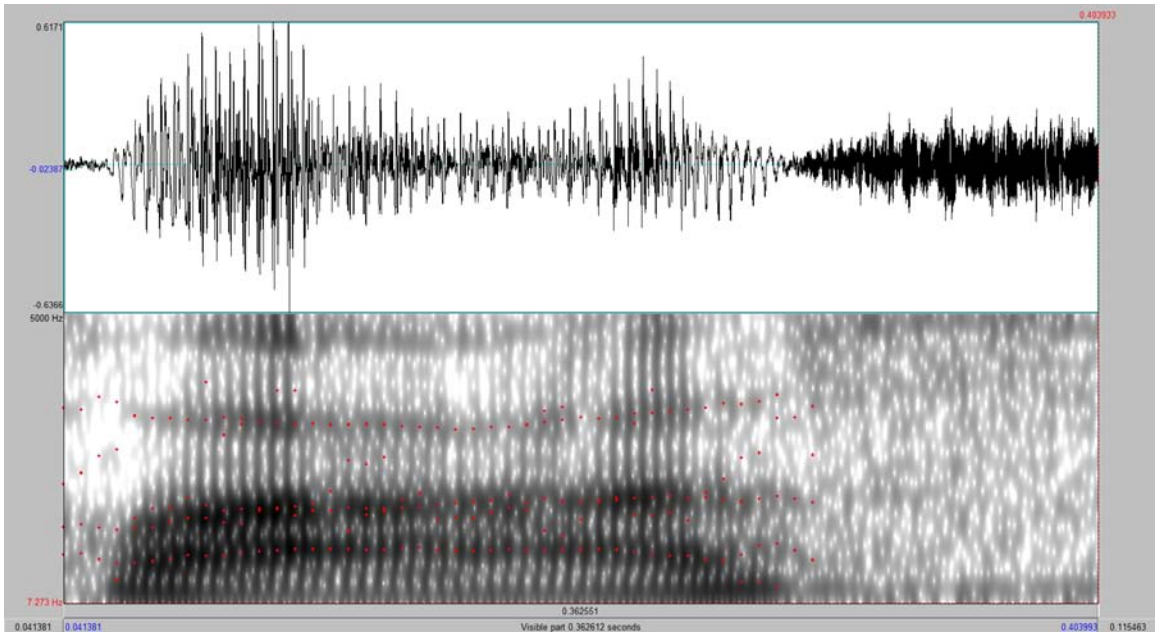

Figure 1: *class* produced by speaker 2 with modal voice (I / 2 / 27:20)

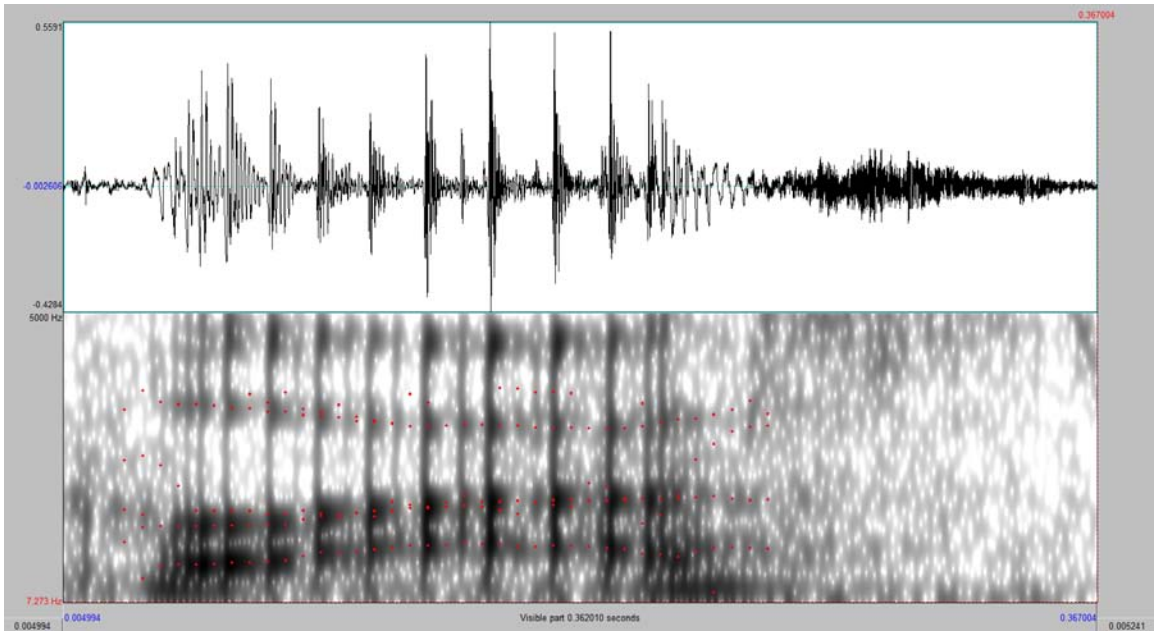

Figure 2: *class* produced by speaker 2 with creaky voice (I / 2 / 28:01)

The differences between these two spectrograms are dramatic, although all program settings (such as window length, pre-emphasis, etc), as well as the zoom level (approximately 362 ms are visible), are identical in the two spectrograms. Comparing the two spectrograms, there are noticeably fewer pulses in the creaky voice example, and their spacing does not appear regular, whereas the modal voice example features a great deal more energy throughout, and a regular, rapid rate of glottal vibration. Another pair of spectrograms demonstrating the contrast between modal and creaky voice follows:

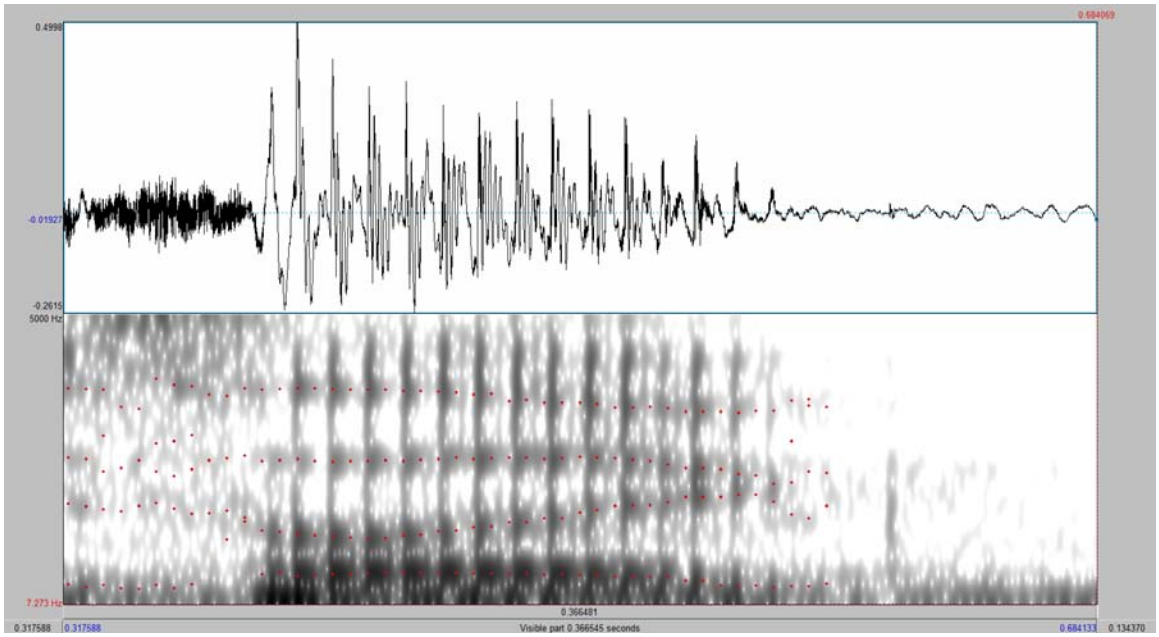

Figure 3: *like* produced by speaker 3 with modal voice (II / 3 / 42:00)

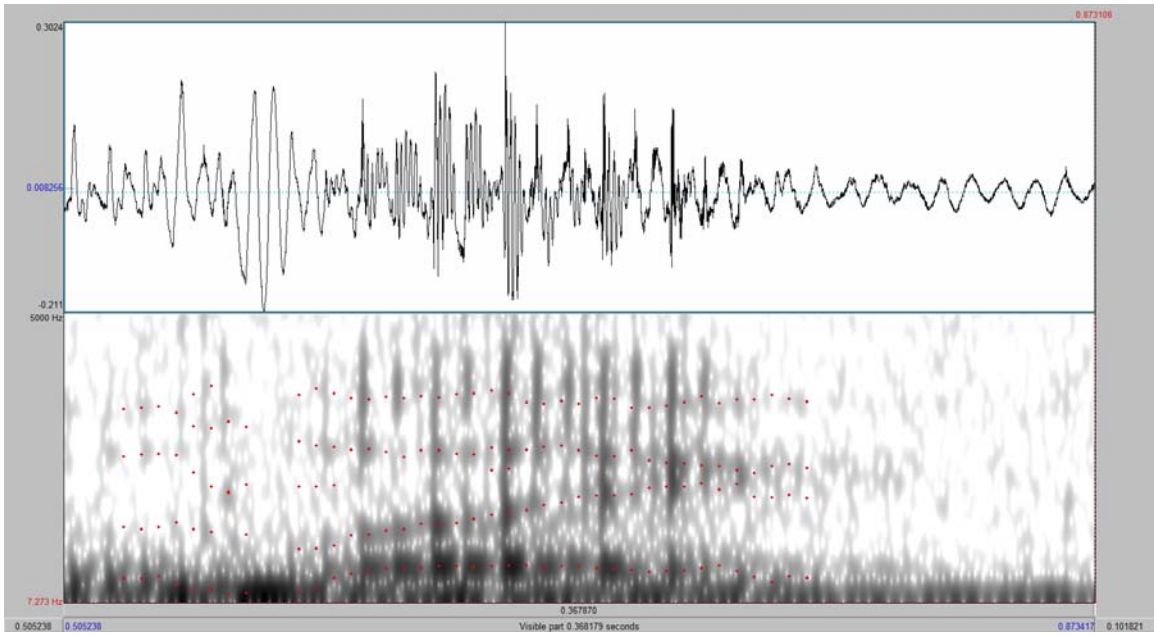

Figure 4: *like* produced by speaker 3 with creaky voice (II / 3 / 3:13)

Again, window length (~36 ms) and all settings are the same, and although the differences between these two spectrograms are not as dramatic as between the first pair,

there are nonetheless some clear indicators that the second spectrogram is creaky voiced. First, there is somewhat less energy present, but most noticeable is the faintness of some of the glottal pulses, with some being relatively strong, and others being rather weak. The vibrations also appear to be slightly irregular, especially when compared with figure 3, wherein the pulses are clearly very regular. In this way, tokens extracted from the interview were spot-checked to determine whether they should be considered creaky or not.

4.1 Speakers

The transcription of the interviews yielded 455 tokens of creaky voice, the breakdown of which is shown in the following table. (I repeat table 1 with the addition of the number of tokens used by each speaker.)

Interview	Speaker	Age	Sex	# of Tokens
I	1	24	Male	117
	2	22	Female	233
II	3	23	Male	89
	4	24	Male	16

Table 2: Speakers and Token Counts

There are two surprising results noticeable in this table: first, over half of the tokens (52%) come from speaker 2, and second, speaker 4 has far fewer tokens than the other three speakers. Regarding speaker 2, this may support the assertion made by some researchers that creaky voice in the Pacific Northwest is more common among women (Ingle et al. 2005), but unfortunately it is not possible to confirm or deny this assertion on the basis of my data, due to the lack of additional female speakers. On the other hand, the

amount of tokens produced by speaker 2 may simply indicate that she was dominant in the conversation, speaking more often than informant 1. Speaker 2's apparently frequent use of creaky voice could also be characteristic of her idiolect, because, as Henton & Bladon have noted, some people are simply "persistent creakers" (1988). Likewise for speaker 4, the low amount of tokens attributed to him may only mean he spoke less than the other participants, however it may also be an indication that speaker 4 has a slightly different phonology from that of the other speakers.

Utilizing the word count generated by ELAN, it is possible to calculate the percentage of each participant's speech which had creaky voice. The following table illustrates these values:

Interview	Speaker	Age	Sex	Total # of Words	# with Creaky Voice	%
I	1	24	Male	5132	269	5.24
	2	22	Female	5644	507	8.98
II	3	23	Male	7041	252	3.57
	4	24	Male	3843	54	1.4

Table 3: Speakers and Word Counts

There are clearly some large differences between each speaker's percentage creaky voice, however these values may not actually be significant, as a linguistically conditioned feature would not necessarily be expected a priori to occur with one any particular degree of frequency, rather the frequency would be determined by how often the trigger comes up. ELAN also generates further statistics from the interviews, including how much time each participant spoke, and what percentage of the interview their speech accounts for.

Interview	Speaker	Age	Sex	Time Spoken	% of Interview
I	1	24	Male	21:52	36.24
	2	22	Female	21:40	35.89
II	3	23	Male	36:09	59.44
	4	24	Male	15:05	24.82

Table 4: Speakers and Time Spoken

Interestingly, these data both confirm and deny earlier hypotheses made about the dominant speaker in each interview. Speakers 1 and 2 spoke for almost exactly the same amount of time, meaning that rather than being dominant in the conversation, speaker 2 spoke either at a more rapid rate, or with fewer pauses. Looking at the data for speaker 3 and 4 however, we have further confirmation that speaker 3 was dominant in their interaction. Speaker 3 spoke more than twice as long as speaker 4, which although significant, is not enough to account for speaker 4's extremely low rate of creaky voice usage.

4.2 Tokens

Analysis of the tokens shows that after creaky voicing is initiated, it tends to remain, and is not typically confined to a single word. The average length of a stretch of creaky voice is 2.38 words, with many longer tokens of 5+ words. The example below shows a longer multi-word token, taken from speaker 2's discussion of the merits of language-learning software.

> "which I guess you could s- argue is like... what it would be like going to ~class but I'd argue that's not the case... I don't know~...
> **Example 2: Multi-Word Token (I / 2 / 33:16.110 - 33:22.610)**

In the data, creaky voice can occur almost everywhere, including various positions within a clause or an utterance (-initially, -medially, and -finally, as well as whole utterances and clauses). In addition, it shows up on discourse markers, fillers, adjacency pairs, and sometimes in transition relevance places. Tokens were coded for their position within a clause as per the examples below, thus, if the first word in a token was the first word in a clause, it was deemed clause-initial, and if the last word in the token was clause-final, it was classified a clause-final token. Similarly, if a token began and ended exclusive of the first and last words in the clause, it was coded as clause-medial. Those tokens whose first and last words were clause-initial and –final respectively were categorized as whole clause tokens, and those that cross clause boundaries were coded as multiple clause. There were also a small number of tokens which consisted only of a filler, and since it is not clear that fillers can be considered part of a clause per se, these tokens were not coded for clause position.

> "and… anyway this guy emails us about… ~he's like~… 'why are you letting destination resorts like Tetherow… build out near Redmond?' and I was like… 'alright first of all…'"
> **Example 3: Clause-Initial Token (II / 3 / 42:36.550 - 42:49.590)**

> "they did all… three of- all of those last year, they went to California, Hawai'i, uh… Glacier National Park, and then they ~did~ some other stuff"
> **Example 4: Clause-Medial Token (I / 1 / 23:36.740 - 23:44.780)**

> "you didn't threaten to punch me in the face I was just afraid you were going to punch me in the face because you seemed very protective of the tire swing and I wanted to play on it, and I had just gotten my face smashed in... when I fell off a swing... and I was just afraid that you might… want to punch me in the nose… I don't know ~why~"
> **Example 5: Clause-Final Token (I / 2 / 1:53.570 - 2:09.230)**

Speaker 3: “well... to be a news anchor in Bend Oregon you really don't... I mean... it's not to...”
Speaker 4: “~it's not like Portland obviously~”
Speaker 3: “~exactly~, it's... market one ninety two, and... out of, I think, two sixteen”
Example 6: Whole Clause Token (II / 27:50.0 - 27:51.470)

Speaker 4: “why was it a good story?”
Speaker 3: “I don't know… ~I thought it was hilarious~”
Example 7: Multiple Clause Token (II / 53:36.160 - 53:39.505)

Similar to clause position, tokens were coded for utterance position by taking into account the beginnings and ends of each token. Depending on the edge(s) the token coincided with, it was classified as utterance-initial, -medial, -final, and whole. Multiple-utterance tokens should not be possible, as the end of an utterance implies a complete termination of vocal activity, and therefore, adjacent utterances must be distinct tokens, rather than the continuation of creaky voice across periods of silence.

Speaker 3: “well... to be a news anchor in Bend Oregon you really don't... I mean... it's not to...”
Speaker 4: “~it's not like Portland obviously~”
Speaker 3: “~exactly~, it's... market one ninety two, and... out of, I think, two sixteen”
Example 8: Utterance-Initial (II / 27:50.0 - 27:51.470)

“and… I… came over and I sat down next to you and I started reading the paper, and then you woke up… and then you were like ‘oh ~hi~’ and I was like… ‘well hello’”
Example 9: Utterance-Medial (I / 1 / 03:58.435 - 04:06.895)

Speaker 2: “and then in the morning we would try to get up and get ready… as fast as we could… to go out and use ~a public restroom~”
Speaker 1: “all of you?”
Example 10: Utterance-Final Token (I / 54:08.320 - 54:15.650)

Speaker 3: “classic high schooler... gossip, right?”
Speaker 4: “~yeah, yeah~”
Speaker 3: “ah it came back to... to haunt me for probably, like four months”
Example 11: Whole Utterance Token (II / 05:57.822 - 06:05.270)

Turn position was also coded similarly to clause and utterance position, although in categorizing tokens for their position within a turn-at-talk, it was necessary to look closely at the discourse, including adjacent utterances from both speakers. The beginnings of turns were defined as utterances following a period of silence or a length of speech from the other participant (including adjacency pairs, which elicit responses from other speakers). Turn-medial tokens were those which occurred clause- or utterance-medially, or which were preceded and followed by utterances from the same speaker.

Speaker 4: “well yeah but… they don’t… not supported… really”
Speaker 3: “~no they aren’t~, nobody makes HD-DVDs already, so…”
Example 12: Turn-Initial Token (II / 51:13.020 - 51:18.970)

“and, um… and then- then- that, that way if they got ~caught~… you know… if uh… the police caught them and it was at their house then…”
Example 13: Turn-Medial Token (II / 4 / 34:31.450 - 34:37.850)

Speaker 1: “I love you and I love that you love these ~stories~ but I ~hate the stories~”
Speaker 2: “well I think they're kind of stupid honestly but…”
Example 14: Turn-Final Token (I / 17:42.180 - 17:47.490)

Speaker 1: “tell me about that ring you have”
Speaker 2: “~uh… I have a ring~”
Speaker 1: “tell me about your teeth you hav- you have a really interesting teeth story”
Example 15: Whole Turn Token (I / 8:14.780 - 8:27.910)

Adjacency pairs were defined in view of Schegloff & Sacks' "Opening up Closing" (1973), which describes them as two-utterance adjacently-positioned pairs whose components are uttered by different speakers. For the purposes of this paper, a token was coded as being part of an adjacency pair as long as one element of the pair was contained within a stretch of creaky voice. The following example shows creaky voice within both parts of typical interrogative, however other adjacency pairs such as pre-closings, (dis)agreements, claims, etc, are also present in the data.

Speaker 1: "so then... uh we were at the, uh, renaissance fair when I ask... well I had asked you originally but then we, like, talked about it some more… and I... thought..."
Speaker 2: "when did we hang out ~before that~?"
Speaker 1: "~we didn't~"
Example 16: Token within an Adjacency Pair (I / 2 / 6:51.130 - 6:52.490)

Tokens were also coded based on whether or not they contained a discourse marker. The identification of discourse markers is in keeping with prior findings on the subject, categorizing words such as *and*, *or*, *but*, *you know*, *well*, *so*, and others as discourse markers (Fraser 1998, Schiffrin 1987). It has been suggested in the literature on discourse markers that saliency allows these discourse markers to achieve the desired effect (De Rooij 2000), and perhaps creaky voice is a way of increasing this saliency. The example below illustrates the use of a well known discourse marker prompting listener involvement (*you know*) with creaky voice, but other discourse markers such as *anyways*, *because*, and *like*, as well as those listed above, occurred in the data.

> "and then I said 'oh hey… [speaker 2], right?' or something like that and then we started talking and you saw my sketches and I was like 'this is embarassing'… ~because you know my sketches were pretty terrible probably~"
> **Example 17: Token with Discourse Marker (I / 1 / 03:18.620 - 03:30.500)**

The last factor that the tokens were coded for was the presence of a filler. Fillers are sounds whose function is to "fill in" the periods of silence between utterances or turns, or when the speaker has to pause to mull something over before continuing. In English, the class of fillers consists of the sounds typically represented orthographically as *uh* and *um*, along with other disfluencies of speech, such as, orthographically, *hm* and *er*. Interestingly, although fillers do not constitute a significant percentage of the tokens, there does appear to be a tendency for the speaker to slip into creaky voice before or after momentary lapses in speech, such as in the following example:

> "there's like some stupid convoluted story about how [name] was the very first person... ~um... that, uh~... I worked with... that I work with... that moved to Bend, she's the first person... piece of talent that was signed to our station"
> **Example 18: Token with Filler (II / 3 / 26:01.430 - 26:19.710)**

4.3 Distributions

Naturally, the tokens are not spread uniformly across the aforementioned environments (see Appendix A for the complete tables): creaky occurs very frequently in some environments, and very rarely in others. Statistical analysis of the tokens shows that the vast majority (63.81%) are in clause-final position, with other clause positions accounting for only small portions of the tokens. Evidence that this environment is

significant comes from the fact that inter-speaker totals are largely the same: the speakers' (1, 2, 3, and 4, respectively) totals for the clause-final environment were 61.82%, 63.48, 69.33%, and 56.25%. The uniformity of these totals across speakers (see Appendix A: Table I) lends strong support to the assertion that this factor is significant in conditioning the occurrence of creaky voice.

A large portion of the tokens are also utterance-final (56.48%), although it is not clear that utterance position in this data represents a relevant factor (see Appendix A: Table II for the complete statistics). As was mentioned in section 3.4, the definition of "utterance" is tenuous, and may not be relevant in this data. Additionally, there is a high degree of correspondence between clause-final and utterance-final, and comparison of the two factors shows that 77.81% of the tokens marked as clause-final are also utterance-final. This high degree of correspondence, added to the fact that clause-final accounts for a greater percentage of the tokens, suggests that utterance position may only account for a significant amount of the tokens due to the tendency for utterance and clause positions to coincide, rather than being causally linked to creaky voice.

The tokens were also coded for their position within turns-at-talk (see Appendix A: Table III), although this factor appears to be irrelevant to the use of creaky voice in the Pacific Northwest, as 58.24% of the tokens occur turn-medially, and it seems unlikely that creaky voice would be conditioned by an environment such as this. Nevertheless, the results for the adjacency pair factor are taken into account, another interpretation surfaces. Only a very small number of the tokens (4.18%) occurred within an adjacency pair (Appendix A: Table IV), which means that 95.82% of the tokens did not occur

within an adjacency pair, that is, they occurred during longer stretches of speech or narrative.

The other factors under analysis were discourse markers (Appendix A: Table V) and fillers (Appendix A: Table VI), however there is little evidence to suggest that these factors play a role in conditioning creaky voice, as discourse markers and fillers account for only 10.99% and 7.25% of the tokens, respectively. Regarding fillers in particular, even when the number of tokens which include a filler is compared against the total number of fillers in the interview, the percentage of the fillers that have creaky voice is still relatively low, being between 8.51-9.28% for speakers 1, 2, and 3 (speaker 4 produced no fillers with creaky voice). Regarding discourse markers, although it would be quite elegant to suggest that creaky voice was being used to increase saliency, a hypothesis along those lines finds little support in these data. Speaker totals for the percentage of discourse markers produced with creaky voice is very low, averaging 1.79% across all speakers, which, needless to say, is far too low to be of any significance.

Additional factors were also considered for the analysis, including topic, lexical category, and quotatives, however all of these were ultimately rejected. Topic appears to be irrelevant to the use of creaky voice because creak is used quite consistently throughout both interviews, despite the fact that the topic changes frequently and sometimes drastically. Lexical category is also unlikely to be significant, as even a brief glance at the data reveals that all classes are regularly found with creaky voice. Lastly, quotatives were excluded from the final analysis after preliminary coding revealed only 11 tokens featuring creaky voice and a quotative.

When speaker token counts were presented earlier on in this paper, I made the observation that there is a large discrepancy between the number of tokens produced by speaker 4, and the number produced by speakers 1, 2, and 3. If creak is phonologically-conditioned, it could be an indication that speaker 4 is a speaker of a different dialect. If creak is physiologically-conditioned, this would be an unexpected result, since, as a male, speaker 4 would not be expected to have a significantly different physiology from speakers 1 and 3. At this point, then, it would be prudent to consider whether or not the data from speaker 4 supports the hypothesis that, having spent all his life in Oregon after the age of 5, he is a speaker of the same variety as my other three consultants. Although speaker 4 clearly talked less than speaker 3 did, it isn't necessary the case that the factors governing the other three speakers' use of creaky voice do not apply to speaker 4. First of all, with the exception of a few protracted stories near the end of the interview, the majority of speaker 4's utterances are responses to speaker 3's stories and questions, and are therefore shorter in length. It is possible then that short utterances reduce the likelihood that creaky voice to be used, however this is not certain. Furthermore, the distributions of the tokens in speaker 4's data are very similar to those of the other participants, which is consistent both with phonological and explanations for creak. This, combined with the fact that speaker 4's vowels patterned with the other three speakers in a previous study, leads me to conclude that despite the small amount of tokens he produced, speaker 4's data should be included along with data from the other three speakers.

5. Discussion

The data presented above do not offer a clear-cut explanation for the use of creaky voice in this variety of English, however there are a number of different hypotheses which can be made about the results. It seems clear that creaky voice is being used far more often by these speakers than physiological factors alone can account for, as physiology would predict creaky voice to occur primarily utterance-finally in English, a position typically marked by falling intonation (Henton & Bladon 1988, Podesva 2007). Although the utterance-final environment is a good predictor of creaky voice in this data, I have shown that clause-final is a better generalization, not only because it captures more of the data, but because the data it captures cannot be adequately explained by physiological factors alone. The following two excerpts are examples of the sort of data which are utterance-medial, but crucially also clause-final:

> "and I didn't know my grandma had some false teeth… ever, until like two years ago, and she 's had them her whole life… well not her ~whole life~, that 's a bit of an exag- maybe like most of my life"
> **Example 19: Clause-Final, Utterance-Medial Token (I / 2 / 09:39.030 - 09:51.450)**

> "I... my mind could not remain focused on the task at hand it was really pretty ~crazy~, I don't know ~how to explain it~ but I 'm sure… people do know what I'm talking about"
> **Example 20: Clause-Final, Utterance-Medial Tokens (I / 1 / 39:51.670 - 40:01.640)**

Both of these examples (particularly the second, where there are two tokens of clause-final creaky voice in a single utterance) are utterance-medial, an environment which accounts for a comparatively small percentage of the data, and one which is a somewhat

implausible environment to trigger a different phonation. Motivating the hypothesis that clause-final position triggers creak is difficult, however there are some reasons to consider it. One possibility is that the phonology of Oregon English may be such that normal drops in pitch at the ends of clauses and sentences are more dramatic than in other varieties of English. A steeper drop in pitch could lead to a concomitant change in the amount of creaky voice used, as has been suggested in the literature (Henton & Bladon 1988, Laver 1994). Another possibility is that creaky voice is a way of signaling the ends of clauses, similar to the way short pauses can be used to set off a subordinate clause.

Another possibility which falls out from the data is that creak is actually being used to mark longer turns-at-talk, signaling the conversant's intention to continue speaking. The fact that the majority of tokens occurred turn-medially and outside of an adjacency pair would seem to support this assertion, as adjacency pairs (which constitute turn boundaries), as well as turn boundaries more generally (e.g. turn-initial and –final environments) are minimally represented in the tokens. Under this hypothesis, creaky voice could be deployed in extended turns-at-talk to signal that the speaker wishes to continue speaking, despite a slower rate of speech or the initiation of word-search, which might otherwise be invitations for another speaker to begin his or her own turn.

The previous two hypotheses are also jointly supported by the observation that most tokens of creaky voice occur at the end of a discourse unit (such a clause or an utterance, where pauses could be anticipated) or in extended turns, where interruptions could be anticipated. Under this conception of the use of creaky voice in Oregon English,

creak would serve to alert other speakers to impending pauses, perhaps as a means of signaling that, despite the silence, conversation will continue.

An interpretation related to the observation that creak occurs in extended turns is that creak is due primarily to a tiring of the vocal folds, which tends to take place in longer utterances and turns-at-talk. This is a potential explanation for a large portion of the data, however this hypothesis fails to predict tokens the appearance of discrete utterance- and clause-medial tokens, which begin and end independent from clause and utterance boundaries. This is not accounted if it is assumed that creak is due to a tiring of the vocal folds, as one would not expect speakers to repeatedly go from energetic to tired in the course of a single utterance, rather it would be expected that the longer the utterance, the more tired, and thus the more creaky, the speech. This is actually observed in the data (see examples 8, 9, 19, and 20), whereas it is seldom the case that token length increases with utterance length.

Suppositions made by other researches (cf. Laver 1994) that creaky voice is used to express emotional affect, whether it be “bored resignation” or simply a lack of confidence, are improbable when applied to the data from this study, and as Henton & Bladon (1988) show (see section 2.2.6) are unlikely to be true even for British English, the variety which was subject to the original claim. Another claim that creaky voice is used to communicate paralinguistic meaning comes from one of my consultants (speaker 1), who has since informed me that he believes he uses creaky voice when he wants downplay a topic that might otherwise be viewed as very important. Although this is a very interesting idea, I do not believe it can account for all the occurrences of creaky

voice in the data, as there are simply far too many tokens, scattered across too many different topics and environments (to say nothing of the potential problems with self-reporting). These and other similar explanations of the role of creaky voice cannot be substantiated, as Henton & Bladon (1988) and Gobl & Ní Chasaide (2003) both demonstrate. These explanations, therefore, must be left aside for the moment.

A final supposition to consider is the null hypothesis: perhaps nothing is going on, and all creaky voice is conditioned by physiological factors possibly in concert with (U.S. American) English prosody. This hypothesis seems to be an unlikely candidate to explain the results of this study, as physiological and prosodic factors would constrain the appearance of creaky voice to a rather narrow set of environments. Creaky voice would be predicted in utterance-final position, possibly around glottal segments, in protracted low-pitched stretches of speech, and in extended utterances due to declination[5]. Firstly, examples such as 8 and 16, wherein two speakers use creaky voice in adjacency utterances, are not consistent with the null hypothesis. Examples like these are very rare, however if they are not purely coincidental, this would suggest a stylistic accommodation on the part of the speakers. Utterance-finality has already been discussed, and glottal segments do not appear to be any more well respresented in the data than other segments, and although I have no empirical data to support this assertion, after having listened to the interviews many times, it is my impression that there are few low-pitched, muttered stretches of conversation that could trigger creaky voice. The vast majority of the speech during the interviews was at volume typical of an indoor conversation, and the

[5] The drop in F_0 observed to occur throughout the course of English utterances (Sternberg, Wright, Knoll, & Monsell 1980)

participants were in general quite animated. As for declination, although this would seem to be an explanation for the earlier finding that creak tends to occur in extended turns, this cannot be the case. As with the hypothesis that creaky voice is a tiring of the vocal folds, this assertion is not supported by the data. Declination predicts a steady increase in creak towards the ends of utterances, and one would not expect to see discrete tokens of creaky voice utterance-medially unless they continued through to the end of the utterance itself, as the continual drop in pitch would make it more likely that speakers would slip into creaky voice, and less likely that they would start and stop.

While intriguing, these findings are far from definitive. The fact remains that this is a small project, based upon data from only four speakers, and though patterns are beginning to emerge, a larger, more comprehensive study is needed in order to capture the bigger picture of Oregon phonology. These findings bear some similarity to those of Ogden (2001) and of Henton & Bladon (1998). Henton & Bladon (1998) found that utterance-final position strongly correlated with the occurrence of creak, concluding that it was reasonable to suppose that creak was functioning as a pre-pausal demarcation. Ogden (2001) found that Finnish speakers tended to use creak at transition relevance places, an environment which is necessarily utterance-final, and most often clause-final. Similarly, many of the tokens in this data were utterance- and clause-final. Where these results differ from previous findings is in the fact that clause-final appears to be a better predictor of creak than utterance-final. Even though utterance-final accounts for a significant portion of the data, clause-final is a better generalization in that it accounts for

data that utterance-final does not: 37 utterance-medial tokens (69.81% of the total number of utterance-medial tokens) which are clause-final.

Research into dialectal discourse such as this contributes positively to the fields of dialectology and conversation analysis, as well as to discourse analysis in general. This research helps to better describe a dialect which has seen little serious scholarly investigation, despite suggestions from many that the "third dialect" is not as homogenous as it might seem (Ward 2003). Not only do these findings add to the growing body of literature surrounding English in the Pacific Northwest, but they also expand our understanding of the many ways in which phonation can be used in language, and the myriad factors which can condition or constrain its use. Different phonations are often overlooked in languages where they aren't lexically contrastive, and so this study can help draw attention to a previously unknown or overlooked recess of language. Naturally though, this isn't the end of the story. This project represents merely the first in-road into the study of phonation in the Northwest, and it is to be hoped that more follow, in order to further our knowledge of the dialect, and of phonation in general.

Transcription Conventions

-	self-interruption
,	short pause
.	pause
…	longer pause, trailing off
?	question intonation
‘’	reported speech
~~	creaky voice

References

Carlson, B. F., & Esling, J. H. (2000). Spokane. *Journal of the International Phonetic Association*, 30 (1-2), 97-102.

Carlson, B. F., Esling, J. H., & Fraser, K. (2001). Nuuchahnulth. *Journal of the International Phonetic Association*, 31 (2), 275-279.

Clarke, S, & Melchers, G. (2005). Ingressive Particles across Borders: Gender and Discourse Parallels across the North Atlantic. In M. Filppula, J. Klemola, M. Palander, & E. Penttilä (Eds.) *Dialects Across Borders* (51-72). Amsterdam: John Benjamins.

Conn, J. (2002). *It’s not all Rain and Coffee: An Investigation into the Western Dialect of Portland, Oregon*. Paper presented at New Ways of Analyzing Variation 31, Stanford, California, USA.

Conn, J. (2006). Dialects in the Mist (Portland, OR). In Wolfram, W., & Ward, B., (Eds.), *American Voices: How Dialects Differ from Coast to Coast* (149-155). Malden, MA: Blackwell.

De Rooij, V. (2000). French Discourse Markers in Shaba Swahili Conversations. *International Journal of Bilingualism*, 4, 447-467.

Esling, J. (1978). The Identification of Features of Voice Quality in Social Groups. *Journal of the International Phonetic Association*, 8 (1-2), 18-23.

Foster, D. W., & Hoffman, R. J., (1966). Some Observations on the Vowels of Pacific Northwest English (Seattle Area). *American Speech*, 41 (2), 119-122.

Fraser, B. (1998). Contrastive Discourse Markers in English. In A. H. Juckert & Y. Ziv (Eds.) *Discourse Markers: Descriptions and Theory* (301-326). Philadelphia: John Benjamins.

Gobl, C., & Chasaide, A. N. (1992). Acoustic Characteristics of Voice Quality. *Speech Communication*, 11 (4-5), 481-490.

Gobl, C., & Chasaide, A. N. (2003). The Role of Voice Quality in Communicating Emotion, Mood and Attitude. *Speech Communication*, 40 (1-2), 189-212.

Gobl, C., Bennett, E., & Chasaide, A. N (2002). Expressive Synthesis: How Crucial is Voice Quality?. Proceedings of the IEEE Workshop on Speech Synthesis. Santa Monica, CA. Paper 52: 1-4.

Gordon, M., & Ladefoged, P. (2001). Phonation Types: A Cross-Linguistic Overview. *Journal of Phonetics*, 29 (4), 383-406.

Hartley, L. (1999). A View from the West: Perceptions of U.S. Dialects by Oregon Residents. In D. R. Preston (Ed.), *Handbook of Perceptual Dialectology Volume 1* (315-332). Amsterdam; Philadelphia: John Benjamins Publishing Company.

Henton, C., & Bladon, A. (1988). Creak as a Sociophonetic Marker. In V. Fromkin, L. M. Hyman, & C. N. Li (Eds.), *Language, speech, and mind: Studies in honour of Victoria A. Fromkin* (3-29). London; New York: Routledge.

Ingle, J. K., Wright, R. A., & Wassink, A. B. (2005). *Pacific Northwest Vowels: A Seattle Neighborhood Dialect Study*. Lay Language Paper presented at the 149th meeting of the Acoustical Society of America, Vancouver, British Columbia, Canada. http://www.aip.org/149th/ingle.html

Labov, W. (2006). *The Social Stratification of English in New York City* (2nd ed.). New York: Cambridge University Press.

Labov, W., Ash, S., & Boberg, C. (2006). *Atlas of North American English*. Berlin: Mouton de Gruyter.

Ladefoged, P. (2005). *Vowels and Consonants*. Malden, MA: Blackwell Publishing.

Laver, J. (1994). *Principles of Phonetics*. Cambridge: Cambridge University Press.

Lippi-Green, R. (1997). *English with an Accent: Language, Ideology, and Discrimination in the United States*. New York: Routledge.

Moosmüller, S. (2001). The Influence of Creaky Voice on Formant Frequency Changes. *Forensic Linguistics*, 8 (1), 100-112.

Ogden, R. (2001). Turn Transition, Creak and Glottal Stop in Finnish Talk-in-Interaction. *Journal of the International Phonetic Association*, 31 (1), 139-152.

Pittam, J. (1987). Listeners' Evaluations of Voice Quality in Australian English Speakers. *Language and Speech*, 30 (2), 99-113.

Podesva, R. (2007). Phonation Type as a Stylistic Variable: The Use of Falsetto in Constructing a Persona. *Journal of Sociolinguistics*, 11 (4), 478-504.

Portland Dialect Study (2003). http://www.pds.pdx.edu/

Reed, C. E. (1961). The Pronunciation of English in the Pacific Northwest. *Language*, 37 (4), 559-564.

(1983). Linguistic Backpacking in the Pacific Northwest. *Journal of English Linguistics*, 16, 78-80.

Riebold, J. (2009). *Local Spoken Here: English in Oregon*. Unpublished manuscript.

Schegloff, E, & Sacks, H. (1973). "Opening up Closings". *Semiotica*, 8, 289-327.

Schiffrin, D. (1987). *Discourse Markers*. Cambridge: Cambridge University Press.

Sternberg, S., Wright, C. E., Knoll, R. L., & Monsell, S. (1980). Motor Programs in Rapid Speech: Additional Evidence. In R. A. Cole (Ed.) *Perception and Production of Fluent Speech* (507-534). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Trudgill, P. (2008). Colonial Dialect Contact in the History of European Languages: On the Irrelevance of Identity to New-Dialect Formation. *Language in Society*, 37, 241-280.

U.S. Census Bureau (2005, January 31). *State of Residence in 2000 by State of Birth: 2000*. Retrieved March 27, 2009, from http://www.census.gov/population/cen2000/phc-t38/phc-t38.pdf

(2007, January 22). *First, Second, and Total Responses to the Ancestry Question by Detailed Ancestry Code: 2000*. Retrieved March 27, 2009, from http://www.census.gov/population/socdemo/ancestry/ancestry_q_by_DAC_2000.xls

Vaux, B., & Golder, S. A. (2002). The Dialect Survey. http://www4.uwm.edu/FLL/linguistics/dialect/index.html

Wassink, A. B., & Conn, J. (2009). The Pacific Northwest English Project. http://www.artsci.washington.edu/nwenglish/index.asp

Ward, M. (2003). *Portland Dialect Study: The Fronting of /ow, u, and uw/ in Portland, Oregon*. Retrieved from the Portland Dialect Study website: http://www.pds.pdx.edu/

Appendix A: Data Tables

Speaker	Clause-Initial	Clause-Medial	Clause-Final	Whole Clause	Multiple Clause	Total
1	6 (5.45%)	16 (14.55%)	68 (61.82)	19 (17.27%)	1 (0.91%)	110
2	23 (10%)	22 (9.57%)	146 (63.48%)	26 (11.3%)	13 (5.65%)	230
3	3 (4%)	5 (6.67%)	52 (69.33%)	10 (13.33%)	5 (6.67%)	75
4	1 (6.25%)	1 (6.25%)	9 (56.25%)	3 (18.75%)	2 (12.5%)	16
All	33 (7.66%)	44 (10.21%)	275 (63.81%)	58 (13.46%)	21 (4.87%)	431

Table I: Clause Position

Speaker	Utterance-Initial	Utterance-Medial	Utterance-Final	Whole Utterance	Total
1	11 (9.4%)	17 (14.53%)	62 (52.99%)	27 (23.08%)	117
2	14 (6.01%)	31 (13.3%)	139 (59.66%)	49 (21.03%)	233
3	8 (8.99%)	5 (5.62%)	44 (49.44%)	32 (35.96%)	89
4	0 (0%)	0 (0%)	12 (75%)	4 (25%)	16
All	11 (7.25%)	17 (11.65%)	62 (56.48)	27 (24.62%)	455

Table II: Utterance Position

Speaker	Turn-Initial	Turn-Medial	Turn-Final	Whole Turn	Total
1	4 (3.42%)	76 (64.96%)	32 (27.35%)	5 (4.27%)	117
2	5 (2.15%)	172 (73.82%)	52 (22.32%)	4 (1.72%)	233
3	43 (48.31%)	8 (8.99%)	35 (39.33%)	3 (3.37%)	89
4	0 (0%)	9 (56.25%)	5 (31.25%)	2 (12.50%)	16
All	4 (11.43%)	76 (58.24%)	32 (27.25)	5 (3.08%)	455

Table III: Turn Position

Speaker	Adjacency Pair
1	6 (5.13%)
2	6 (2.58%)
3	6 (6.74%)
4	1 (6.25%)
All	19 (4.18%)

Table IV: Adjacency Pairs

Speaker	DMs (Creaky)	DMs (Total)
1	7 (5.98%)	646 (1.08%)
2	24 (10.3%)	683 (3.51%)
3	15 (16.85%)	888 (1.68%)
4	4 (25%)	561 (0.71%)
All	50 (10.99%)	2778 (1.79%)

Table V: Discourse Markers

Speaker	Fillers (Creaky)	Fillers (Total)
1	9 (7.69%)	97 (9.28%)
2	4 (1.72%)	47 (8.51%)
3	20 (22.47%)	221 (9.05%)
4	0 (0%)	68 (0%)
All	33 (7.25%)	433 (7.62%)

Table VI: Fillers