‘I Am a Perpetual Underdog’: Lady Gaga’s Use of Creaky Voice in the Construction of a Sincere Pop Star Persona

Lewis G. Esposito

An honors thesis submitted in partial fulfillment of the requirements for
the degree of Bachelor of Arts in Linguistics
Swarthmore College
December 2015

Abstract

Creaky voice, a vibratory phonation produced when the cartilage from the front and back of the larynx are pressed closely together (Zimmerman 2013), has received significant media attention in recent years, particularly for its prevalence in the speech of female American pop stars, such as Kesha, Britney Spears, and Lady Gaga. Recent studies (Yuasa 2010; Podesva 2013) have found that the phonation is more pervasively employed by American women than men, which is in contrast to earlier studies showing its associations with upper-class males in the UK (Henton-Bladon 1988; Esling 1978; Stuart-Smith 1999). Yet, until recently (Zimmerman 2015; Levon 2015), few studies have extensively analyzed its potential social meanings from an intraspeaker perspective. Additionally, none examine its usage among the women noted above. Firmly situated within the realm of third-wave intraspeaker variation studies, this thesis considers Lady Gaga’s use of creaky voice across four different speech samples -- two interviews and two speeches – to discern her possible stylistic usage of the phenomenon. I find that the calculated percentage of creak is highest in the most personal of the samples and lowest in one of the speeches. Considering these results along with the conversational contexts in which creak is most prevalent, I argue that she employs creaky voice to take an affective stance of intimacy and convey character traits of openness and sincerity. Further, I argue that these meanings arise out of creak’s iconicity with a ‘contained’ emotional state. This thesis (1) adds to the growing body of literature examining the indexical meanings of creaky voice in American English, (2) offers evidence for stance-based persona construction, and (3) emphasizes the importance of considering variation in interaction.

*I would like to thank my advisor, Nathan Sanders, for his unwavering support throughout this process. I would also like to thank Kara Becker and Sameer ud Dowla Khan for allowing me serve as an RA on their Language and Gender project, which inspired this thesis. I am also grateful to Laura Katz and Z.L. Zhou, and Jamie Thomas for their advice on drafts of this paper.
# Table of Contents

1 Introduction .......................................................... 4

2 Developments in the study of variation ...................... 5
   2.1 The first wave: Broad correlations ...................... 5
   2.2 The second wave: Variation at the local level .......... 6
   2.3 The third wave: Stylistic practice and persona construction .... 7

3 Style and social meaning .................................. 8
   3.1 Style in the third wave ................................ 9
   3.2 Social meaning and indexicality ..................... 9
   3.3 The social meaning of creaky voice .................. 11

4 Lady Gaga ...................................................... 12

5 Methods and Study Design ................................ 12
   5.1 Speech samples .......................................... 12
   5.2 Creak identification ................................... 14
   5.3 Parenthetical, intonational phrase, and glottalization identification .... 16

6 Results .......................................................... 17
   6.1 Structural creaky voice ................................ 19
   6.2 Structural and non-structural creaky voice across speech samples .... 22

7 Meaning in Lady Gaga's creaky voice .................... 24
   7.1 Revisiting the Attention to Speech model ............... 24
   7.2 A third-wave approach to meaning ..................... 25
   7.3 Creaky voice: Towards a stance of intimacy .......... 26
   7.4 Intimacy and iconicity .................................. 31
   7.5 Openness and sincerity .................................. 32

8 Beyond Lady Gaga: Indexical field for creaky voice .... 35

9 Conclusion ..................................................... 36
Creaky voice, often referred to as vocal fry in popular media, is a type of non-modal phonation characterized by a low frequency and vibratory sound akin to that of an unoiled door hinge. It is produced when the cartilage from the front and back of the larynx are pressed closely together, causing the vocal cords to grow lax (Zimman 2013). Creaky voice is phonemically contrastive in some languages, such as Zapotec languages (Sicoli 2010) and Northwest Native American languages (Gordon and Ladefoged 2011).

Recently deemed trendy among "upwardly-mobile" American women (Yuasa 2010), creaky voice is famously present in the speech of many female pop-cultural icons. Britney Spears, Kesha, Lady Gaga, Katy Perry, and Kim Kardashian are often cited as classic creakers in popular news articles, and they are subsequently blamed for inciting the trend among younger generations. Indeed, the policing of linguistic innovation – particularly as it relates to the speech of young women – is pervasive, relentless, and real. Lexicon Valley podcaster Bob Garfield has called the creaky voice “vulgar” and “repulsive”, while a fine arts professor at Hofstra University points to it as a revelation of “a radical uncertainty about one’s place in the world” (Hess 2013).

While sociophonetic studies on non-modal phonation patterns were at one time scant (Podesva 2013), this recent obsession with creaky voice in the media seems to have seeped into the linguistics literature – or, perhaps, vice versa. Most studies considering creaky voice have correlated it with sex. Three of the earliest and most cited studies, Henton-Bladon 1988, Stuart-Smith 1999, and Esling 1978, found that men in the UK are significantly creakier than their female counterparts; this trend seems reversed in American English: Leftkowitz and Sicoli (2007), Yuasa (2010), and Podesva (2013) have reported that American women are creakier than men. Other studies, such as Zimman 2013 and Becker 2015, have emphasized the importance of considering self-identifying gender and sexuality in the distribution of creak. Some studies have considered creaky voice’s role in discourse (Mendoza-Denton 2011; Podesva 2013; Zimman 2015; Levon 2015) but have come to different conclusions. Intraspeaker analysis of creaky voice is limited to Levon 2015 and Zimman 2015, so this thesis seeks to add to the literature in this domain, which is summarized in sections §3.3.

At its core, third-wave variationism asks what speakers hope to achieve socially through the use of speech variants; the methodologies of this framework are outlined in §2. As previous studies of creaky voice in American English have largely focused on its distribution across the population, this work examines its distribution across different speech contexts...
for a single speaker, as well as its potential social meanings – traits that the speaker and listener associate with the variable. Additionally, considering how often the female celebrities noted above are referenced for their creak usage, it is surprising that their speech has not yet been analyzed rigorously to discern the actual extent of their usage. Filling this gap, I consider the speech of Lady Gaga in four different contexts: (1) a personal, biographical interview; (2) an award acceptance speech; (3) an album discussion interview; and (4) a speech for a gay rights rally. Finding that she uses creaky voice most extensively in the personal interview, and considering the interactional moments in which creak appears, I argue that she uses it to take an affective stance of intimacy; this stance indexes character traits of openness and sincerity. Further, I contend that as she repeatedly takes this stance, she is constructing a sincere pop star persona to appeal to her fans.

Beyond the literature review in sections §2 and §3, this thesis is divided into six additional sections. §4 is a short background on Lady Gaga. §5 discusses my methods of analysis, including sources for the speech samples, creak-discernment techniques, and statistical methods. §6 includes my quantitative findings and discussion. §7 adopts a qualitative approach to discern the social meaning of creaky voice, and §8 considers the social life of creaky voice beyond Lady Gaga. §9 is the conclusion.

2 Developments in the study of variation

Sociolinguistic variation studies have a history originating with William Labov and his 1966 study The Social Stratification of English in New York City. Since then, variationists have developed a variety of different frameworks to consider variation in its social landscape, none of which are impervious to criticism. Eckert (2012) broadly conceptualizes these different methodologies as occurring in three waves – the final of which is the framework adopted in this paper; the waves are discussed in the following sections.

2.1 The first wave: Broad correlations

The first wave of variation studies, propelled by Labov, explain phonetic variation as resulting from conflicting forces within the linguistic system, first reaching those least influenced by the standard variety, and later those most immersed in it (Eckert 2012). Methodologies involve large surveys as well as recorded interviews which correlate variables in speech to macro-demographic categories like socioeconomic class and gender (Eckert 2003). As
Eckert (2012:88) notes, “In this way, speakers emerged as human tokens – bundles of demographic characteristics.”

Prominent in the first wave is the concept of the vernacular, which represents speakers’ most natural way of speaking (Labov 2001). The social meaning of variables is limited to socioeconomic status, and differences in gender and attention paid to speech “were seen as resulting from the effects of these categories on speakers’ orientation to their assigned place in that hierarchy” (Eckert 2012:90).

Criticisms against this approach have been raised. While we would expect those from the lowest socioeconomic groups to engage least with the standard variety, they are, in fact, not the leaders in sound change – rather, those of the upper-working and lower middle classes are the pioneers. As the lower middle classes engage most frequently with both the vernacular and the standard, their high incidences of vernacular variants suggest that they index some type of positive social value (Eckert 2012). Further, it has been found that associating sound change with sex yields misleading results. While women have classically been shown to lead in variation, it was later discovered that the speech of upper-middle-class women was more standard than that of upper-middle-class men, while working-class men’s speech was more standard than working-class women’s (Labov 2001). Eckert and McConnell-Ginet (2013:257) caution against approaches correlating variables with broad demographic categories, “…the statistical generalization[s] hide the more interesting truths, because the significance of these linguistic differences lies in the question, ‘which women and which men and why?’” In any case, first-wave studies are indispensable. They reveal the distribution of large-scale distribution – important data for discerning the mechanics and driving linguistic forces in sound change. They also provide a basis for studies in the following waves, which consider, through ethnographic methods, how these changes play out on a smaller scale.

2.2 The second wave: Variation at the local level

The variation studies of the second wave center on ethnographic research and linguistic change in smaller communities, such as self-identified social and ethnic groups (Eckert 2012). According to Eckert, Milroy’s (1980) study of phonological variation in the social networks of Belfast signaled the beginning of the second wave. Milroy argues that variables correlate more aptly with social identity groups than with macro-level demographic categories; further,
Milroy neglects the view of variation as being passively determined by the greater linguistic system.

Eckert (1989, 2000) conducted ethnographic fieldwork in high schools situated within a largely white Detroit suburb. She discovered two opposing social groups in the high schools: “jocks” and “burnouts”. The preppy jocks centralized their identities in their academics and extracurriculars, while the burnouts pursued a largely vocational curriculum and rooted their social networks in neighborhood life. The burnouts employed more features of the local vernacular than the jocks, and while the former generally belonged to the lower rungs of the socioeconomic hierarchy and the latter to the upper ones, social identity group was a better predictor of phonological variation, such as the backing of /e/ and raising of /aɪ/, than family income; this finding suggests that variation does not crystallize in early childhood development as a result of the class and education of parents, but that it evolves with children as they situate themselves in the social landscape (Eckert 2012). Further, Eckert found that a subset of the burnout girls, the “burned-out burnout girls”, recognized as the most radical amongst the burnouts, had the greatest usage of vernacular variants, including phonological and grammatical variation. Based on this discovery, she argues that variables do not directly index category membership – if they did, we would expect all members of the burnouts to exhibit similar usages of vernacular variants – but characteristics associated with these categories, such as radicalness. Thus, second-wave studies move beyond the macro-sociological categories of the first wave in favor of attributing linguistic variation to social groups identifiable only through ethnographic insights.

2.3 The third wave: Stylistic practice and persona construction

While second-wave variation studies build upon the findings of the first wave through a consideration of micro-level variation, they still link variables to static groups, like sound changes characteristic of the Northern Cities Vowel Shift to the jocks and burnouts noted above. The impetus of the transition from the second to the third wave placed a new emphasis on the agency of speakers’ stylistic variation (Eckert 2012), as well the social meaning of this style (Eckert 2008), and how individual speakers design these styles and imbue them with meaning (Wolfram and Schilling 2015).

Situated at the center of third-wave studies is the pursuit of social meaning, which, in the third wave, refers to the traits that speakers and listeners attribute to a variant, and what exactly speakers are using this variable to achieve. The third wave emphasizes social agency
in variation, implying that speakers can, to certain extent, use variables to construct different speech styles and fleeting personas, and, in the long term, identities. In a sense, then, third wave studies seek answers from an even smaller scale – the individual speaker – and what they do with variation from moment to moment. Podesva (2011), for example, found that his consultant Regan exhibited features of the California Vowel Shift to the greatest extent when on a boy’s night out and least when talking to his supervisor. Podesva argues that the features are part of a stylistic package portraying a partier persona, related to gay identity.

The extent to which speakers can alter their speech is often contested, and some studies in this wave have been criticized for not considering the greater picture of variation and its linguistic constraints (Tagliamonte 2016). Gregory Guy notes in an interview: “You cannot, if you’re a working-class black kid growing up in North Philadelphia, just choose by means of your agency to talk like an upper middle class white person from Chicago...and what you do with your experiences, that’s where you have the agency, but the experience defines the outer limits of what’s possible” (Tagliamonte 2016). Third-wave studies nonetheless are irreplaceable within the domain of sociolinguistic variation studies for their emphasis on the social and what speakers are actually doing with particular speech variants. For this reason, I choose to consider creaky voice from a third-wave framework.

3 Style and social meaning

*Style* refers to the variation in the speech of individual speakers (Wolfram and Schilling 2015). Very early definitions, like that of Labov (1966), hold that style exists within a casualness/formality continuum: the more attention speakers pay to their speech, the more standard it becomes (Labov 2001); similarly, less attention paid to speech results in less standard varieties. In this case, speakers’ agency is limited to their ability to avoid stigma. A later conception by Bell (1984) is known as the Audience Design model. Rooted in Speech Act Theory, it states that people adjust their speech either toward (convergence) or against (divergence) the addressee. Most recently, style in third-wave variation studies highlights speakers’ capacities to alter their speech agentively; this concept is further explored in §3.1

Social meaning, on the other hand, was not necessarily the object of analysis in first and second-wave studies and thus lacked a proper framework of consideration until recently. This third-wave framework is discussed briefly in §3.2. Previous analyses of the social meaning of creaky voice in American English are provided in 3.3.
3.1 Style in the third wave

Third-wave variationists consider style as it unfolds in interaction (Wolfram and Schilling 2015), and this view is akin to the speaker-design approach termed by Schilling-Estes (2002). Styles are composed of any number of variables which have a generalized meaning, often discernible through first-wave variation studies, that become more specific in context (Podesva 2013). Podesva (2007), for example, considers a gay doctor’s use of falsetto. He argues that falsetto has a generalized meaning of expressiveness, but, based on the discourse and situational contexts in which it appears more often, assists in the construction of a diva persona. As Podesva (2007:14) notes, “I argue that the social and linguistic contexts in which a variable is uttered color its social meaning, enabling the variable to participate in the construction of more specific, identity-based meanings.”

Central to style in the third wave is the ability of speakers to employ an aggregate of variables to create new categories, identities, and social meanings. Zhang (2005) presents a prime example of speakers’ constructions of character types that are not widely circulating (stoner, diva, nerd), but are new innovations. She shows that a growing group of Beijing elites (“yuppies”) employed in the global financial market have abandoned two features typical of Beijing Mandarin: word-final rhotacization, often associated with sleaziness, and the lenition of /z/ to an interdental, ideologically linked with lazy character types. Instead, they adopt a full tone where Beijing Mandarin has a neutral tone (shortened and less emphasized). Through the use of variables ideological connected to identity types and traits in the social world, the yuppies construct for themselves a style international in scope and thus in opposition to the locals.

3.2 Social meaning and indexicality

As noted, social meaning was not the object of analysis until third-wave variation studies. Eckert (2008) seeks to develop one with her *indexical field*.

The indexical field is built on Silverstein’s (2003) concept of indexical order and its ability to link broad ideologies with more specific meanings in context. A first order index (variable) marks membership in a community, like the yuppies, jocks, and burnouts cited earlier. These are similar to Labov’s *indicators* in that they point to group membership but are not stigmatized. A second order index refers to the characteristics associated with these categories – the rebelliousness of the burnouts, for example. This indexicality allows for the continued reassignment of value and meaning through the *nth order usage*; this system
Figure 1: Indexical field of hyperarticulated /l/ in American English. Boxes = character types; black = qualities; gray = fleeting stances.

is reflected in Eckert’s (2008) notion of the indexical field. An indexical field, then, is a network of meanings for a variable that are linked through their social ideologies. It is as flexible and dynamic as are the ideologies that are associated with these variables.

As Eckert (2008:464) explains, “Variables have indexical fields rather than fixed meanings because speakers use variables not simply to reflect or re-assert their particular or pre-ordained place on the social map, but to make ideological moves.” This view of variables can thus capture and synthesize their broader ideological meanings with the speaker’s unique adaptations of these meanings. Podesva (2007; 2013) finds that falsetto voice is widely used to stylistically convey expressiveness. When his consultant Heath uses it amongst his friends, it assists in the production of a “diva persona”; in the speech of an African American woman, it is adopted to express authority and power. Variables thus have generalized meanings that often emerge out of associations with the people who use them, as well as iconicity with the variable itself, and the contexts in which they are employed. Figure 1 is an example of an indexical field for hyperarticulated /l/ in American English. The different social meanings are those proposed by sociolinguists who have studied the variable through ethnomethodological work and through perception studies.

Refer to Eckert (2008) for a more in-depth discussion of this model, but broadly, it reveals the meaning of precision generally associated with hyperarticulated /l/, partially as a result of the extra effort and thus emphasis that accompanies its production, as well as its stereotypical association with the British, themselves associated with refinery and education, and thus precise speech. As is evident from this brief discussion, these meanings build
off each other through ideological connections and speakers’ creativity in expanding these connections.

### 3.3 The social meaning of creaky voice

Developments in the social meaning of creak have been varied. Some work has associated creak with males and stereotypically male traits. Henton-Bladon (1988) found that creak was more prevalent among men than women in Edinburgh, and Stuart-Smith (1999) found the same in Glasgow. Carpenter (2006) considered the speech of American teenage males and argued for creak as an index of masculinity. Along a similar vein, Mendoza-Denton (2011) analyzed the gang narratives of a group of gang-affiliated Latinas in Northern California and found that they adopted creaky voice most extensively in these stories, perhaps to construct a tough, hardcore persona. Leftkowitz and Sicoli (2007) have asserted that women adopt creaky voice to convey authoritativeness, perhaps due to its low pitch and thus similarity to the voices of men.

Yuasa (2010), on the other hand, contends that the socio-indexical meanings of creaky voice have diverged from associations with authoritativeness and masculinity; in her perception study, she found that listeners rate creaky voice as sounding more “casual”, “educated”, “genuine”, and “compliant” than the corresponding modal voice. Listeners selected terms such as “graduate student”, “urban”, and “professional” to describe creaky voice, and thus she concludes that women might adopt it in the construction of a professional, upwardly-mobile feminine persona, capable of competing with male peers. Podesva (2013) generally agrees with Yuasa’s (2010) sentiments, but still argues that the general social meaning of creaky voice’s indexical field lies in its similarity to men’s voices.

Still, other meanings of creak have been identified, complicating the issue. Brown and Levinson (1987) have suggested that creak might be used to commiserate or complain. Pennock-Speck (2005) ascertained that female actresses adopt more creaky voice in American films and often pair it with breathy voice in intimate scenes; he proposed that it may index feminine desirability and intimacy. Most recently, Zimman (2015) and Levon (2015) have argued for speakers’ use of creaky voice in signaling affective distance. This thesis seeks to reconcile some of these differences.
4 Lady Gaga

Lady Gaga, American singer, songwriter, and actress, was born Stefani Germanotta on March 28th, 1986, in New York’s Upper East Side. She attended New York City’s prestigious Covenant of the Sacred Heart, although she laments her time there, noting, “I was called really horrible, profane names very loudly in front of huge crowds of people, and my schoolwork suffered at one point... I was so ashamed of who I was” (Kristof 2012). Her traumatic experiences with bullying in high school currently inform her philanthropic efforts today, as she fights for social equity in LGBTQ+ communities. In collaboration with her mother Cynthia Germanotta, she founded the Born This Way Foundation, “dedicated to creating a safe community that helps connect young people with the skills and opportunities to build a kind, braver world.” Recently, the foundation has partnered with Yale’s Center for Emotional Intelligence and with Elton John’s AIDS foundation to offer mental support to young adults. She notes of her reasons for creating the organization, “Teens who take their lives young because they feel different, I wanted to make [their deaths] a lesson instead of a casualty of our negligence” (Dodero 2015).

Paramount to Lady Gaga’s identity and success as a pop star is her admiration of and utter devotion to her fans, whom she endearingly calls “little monsters”. Famous for her eccentric aesthetics, she is constantly re-inventing her image, style, and persona. Of the American pop stars often referenced for their extensive creak usage, I chose her specifically for this tendency.

5 Methods and Study Design

5.1 Speech samples

Four samples of Lady Gaga’s speech were selected for analysis from YouTube. The limitations of solely considering interviews for intraspeaker variation studies is well known, so I selected two speeches and two interviews for more of a variety. The primary criteria used in selecting interviews was length and date recorded – all speech samples were 8-10 minutes in length and were recorded sometime during 2011-2013, the height of Lady Gaga’s musical career. After the videos were collected, their audio tracks were converted to MP3 files via ClipConverter, a free online MP4 to MP3 converter. The url to this website is provided in the appendix, as are those to the videos.
Descriptions of the four speech samples selected are as follows. The first is an interview produced by MTV in 2011, entitled “Inside the Outside”. Less of an interview and more of an autobiographical account of her rise to fame and family background, the camera is focused solely on her – the face of the interviewer is never shown. This sample, content-wise, is the most personal. The second is her acceptance speech for the 2011 Council of Fashion Designers of America (CFDA) Award. This award recognizes fashion trendsetters. Lady Gaga’s speech was more off-the-cuff and conversational than prepared, as she jumped from personal anecdotes, to her collaborators, to the importance of the award to her. She did not read from a script, although she presented in front of a large audience. The third sample is a 2013 one-on-one interview produced by The National, a televised news outlet in Canada. Lady Gaga talks about her album ARTPOP, philanthropic efforts, musical and artistic style, and views on celebrity culture. The final video is a speech given in Rome, Italy, at Europride 2011, a European LGBTQ+ pride celebration hosted by a different city every year. She read entirely from a prepared script, as she addressed LGBTQ+ discrimination issues and the significance of social equity and marriage equality.
5.2 Creak identification

Each sample was transcribed via TextGrids in Praat (Boersma and Weenink 2015), and a script was used to extract the transcription and timestamps into a .txt file. The .txt file was then uploaded into FAVE (Rosenfelder et al. 2011) for auto-alignment. The output was loaded in Praat to code for voice quality. All vowels longer than 0.02 seconds were coded as creaky or non-creaky. Non-creaky vowels in this study were modal, breathy, or falsetto voice. Where necessary, vowel boundaries were adjusted to account for FAVE’s errors. Creak was identified primarily auditorily; other studies have preferenced auditory judgments over visual clues as well (Yuasa 2010; Becker, et al 2015; Podesva 2011; Sinae 2015). If a vowel was not clearly creaky, visual clues were considered. Mendoza-Denton (2011:264) notes that the waveform of a creaky segment “contains both jitter (irregularity in the period of the wave) and shimmer (irregularity in the amplitude”).

Figures 2 and 3 display the waveforms and spectrograms of tokens selected from Lady Gaga’s spontaneous speech. Notice that the creaky spectrogram of Figure 2 has three pulses, while the modal spectrogram of Figure 3 has seven. Also, as expected, the creaky waveform exhibits both jitter and shimmer: the amplitudes of the waves are far less regular in Figure 2 than in Figure 3, and the vertical distance between each wave is also more variable in Figure 2 than in Figure 3. Factors such as these were taken into account to make a more definite judgment of phonation when auditory cues were unclear. In some cases, only a portion of a given vowel could be identified as creaky; these vowels were coded as creaky if at least 0.01 sec of the vowel was creaky. Others were marked as non-creaky.
Figure 2: Creaky /æ/, 0.05 sec.

Figure 3: Modal /æ/, 0.05 sec.
5.3 Parenthetical, intonational phrase, and glottalization identification

Parentheticals are noted for their transience: an argument is interrupted and acontextual information (the parenthetical) is inserted. Structurally, parentheticals are “marked as being ‘different’ in comparison to the surrounding passages by a number of phonetic and syntactic characteristics” (Ruchota 1998:121). In addition to coding for creak, I identified parentheticals in each sample as well, based on five traits outlined in Lee (2015): (1) faster speech rate, (2) lower pitch, (3) falling-rising intonation marking the end of a parenthetical, (4) slower speech marking a return to the previous thought, and (5) interrupted syntax prior to the insertion of the parenthetical. As Lee (2015) argues for a co-occurrence between creaky voice and parentheticals in discourse, it is important to rule out the possibility that increased creak usage across contexts is appearing as a result of an increased usage of parentheticals.

Indeed, not all parentheticals will have each of the five of the identifying characteristics noted above. I thus coded an utterance as a parenthetical if it exhibited at least three. Suspected parentheticals in speech were noted and were then either confirmed or rejected if they had less than three of the typical traits. Speech speed was analyzed perceptually, and average pitch was extracted from Praat (Boersma and Weenink 2015) for comparisons. As there is clearly no uniform drop in pitch to mark a parenthetical, I arbitrarily set a minimum of 20 Hz; that is, suspected phrases with a pitch of 20 Hz or lower than the surrounding utterances were deemed positive for characteristic 2 listed above.

To provide an example of a parenthetical in context, I offer an adapted sample from Lee (2015) in Excerpt 1. Each line in the excerpt represents an intonational phrase (IP). Note that a brief discussion of intonational phrases follows this excerpt.

Excerpt 1: Doggy day care

1. and that’s another whole snotty neighborhood
2. in the Kenmore area
3. but um
4. you know it’s my fault I guess for working there, what can I say but
5. like, people there are assholes.
In this excerpt, the speaker is discussing her day job and opinion of the Kenmore area (a pseudonym) in D.C. Lines 1, 2, and 5 compose the main narrative, while line 4 is an aside— a parenthetical. The speaker foreshadows a coming change in attitude toward Kenmore with ‘but um’ (line 3) (Lee 2015), and this change comes to fruition in line 4. In contrast to the assertions of negativity directed towards Kenmore (‘whole snotty neighborhood’, ‘people there are assholes’), the speaker in line 4 shifts the blame to herself for deciding to work there. Lee notes that this reversal in the evaluation frame is reflective of the parenthetical, as it signals a break away from the principal discussion. While Lee uses discourse analytic techniques as one of the criteria in discerning the parentheticals here, I did not in an attempt to safeguard uniformity and avoid interpretive judgments.

In addition to parentheticals, each speech sample was broken down entirely into IPs, as previous work has shown that creaky voice favors the IP-final position (Henton-Bladon 1988; Yuasa 2010; Podesva 2007, 2013; Lee 2015). An IP is a prosodic unit of speech containing at least one syllable that receives phrasal stress. They end with specific tone boundary contour: high-high, low-low, high-low, or low-high (Sameer ud Dowla Khan, personal communication). Each boundary contour generally accompanies a type of utterance; for example, high-high tone often co-occurs with up-talk and yes-no questions, and low-low accompanies declarative sentences. In this paper, the actual contour of each IP identified is irrelevant; they were solely considered to discern the end of an IP.

Finally, all creaky vowels were screened for glottalized segments within the same syllable, as glottalization favors the appearance of creaky voice (Podesva 2007; Levon 2015). Glottalization in this study was limited to word-initial vowels, where a glottal stop might precede the vowel, and word-final /t/ and /p/. Given that there were no examples of creaky voice occurring in a syllable with word-final /p/, only word-final /t/ is relevant in this study.

6 Results

Across all four samples, I coded a total of 5,008 vowels, of which 862 (17.21%) were creaky. Lady Gaga employs the most creaky voice in the “Inside the Outside” interview (29.92%), relatively equal amounts in the CFDA Fashion Icon Award acceptance speech (16%) and The National interview (16.54%), and least in the Europride speech (1.5%). These results are all evident in Figure 4 and Table 1; note that the “Inside the Outside” interview is abbreviated Ins., the CFDA acceptance speech is abbreviated CFDA, The National interview is abbreviated Nat., and the Europride speech is abbreviated Eur. in both of these.
Table 1: Total creak usage across speech samples

<table>
<thead>
<tr>
<th></th>
<th>Ins.</th>
<th>CFDA</th>
<th>Nat.</th>
<th>Eur.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creaky vowels</td>
<td>451</td>
<td>163</td>
<td>232</td>
<td>16</td>
</tr>
<tr>
<td>Non-creaky vowels</td>
<td>1056</td>
<td>856</td>
<td>1171</td>
<td>1063</td>
</tr>
<tr>
<td>Total vowels</td>
<td>1507</td>
<td>1019</td>
<td>1403</td>
<td>1079</td>
</tr>
<tr>
<td>Percent creaky vowels (%)</td>
<td>29.92</td>
<td>16.00</td>
<td>16.54</td>
<td>1.50</td>
</tr>
</tbody>
</table>

A Pearson's Chi-Squared test in R (R Core Team 2015) was used to determine possible independence between the two variables tested: speech sample and the proportion of creaky vowels to non-creaky vowels. The contingency table consisted of four columns, one representing each speech sample, and the two rows contained raw numbers for creaky vowels and non-creaky vowels in each sample. I found statistical significance ($\chi^2 = 359.82$, df = 3, p-value < 0.0001), suggesting that the differences in creaky voice usage across samples is not a coincidence.
6.1 Structural creaky voice

After calculating the total percentage of creaky vowels in each sample (Table 1), I sought to determine whether the observed differences were due to differing linguistic and structural conditions across samples. Note that I call creaky voice potentially influenced by one of these factors as structural creaky voice, and creaky voice appearing elsewhere as non-structural creaky voice. The first factor considered was the IP-final position. The IP-final position favors creaky voice as the voice trails off, resulting in a lower f0 (Henton-Bladon 1988; Yuasa 2010; Podesva 2007, 2011; Lee 2015). Additionally, D’Onofrio et al. (2013) have noted that non-final creaky voice has the most potential to be adopted stylistically. I coded creaky voice as IP-final if the final syllable in an IP was creaky. I found that approximately the same proportion of creaky voice in each sample appeared in the IP-final position ($\chi^2 = 1.211$, df = 3, p-value = 0.75) (Figure 5).
Figure 5: Percentage of creaky vowels appearing in the IP-final position out of total creaky vowels

The second structural factor considered was parentheticals. As discussed previously, Lee (2015) has shown that parentheticals and creaky voice often co-occur. The breakdown is evident in Figure 6, and speech sample appears to be the cause of the differences ($\chi^2 = 41.28$, df = 3, p-value < 0.0001), where the CFDA has the largest percentage of creak present in parentheticals while the other samples are not significant from each other (p = 0.38).

In contrast to Lee (2015), who found that 36% of the creak in her phrases could be attributed to parentheticals, far fewer instances of creaky voice in my samples could be associated with them – 3.33%, 14.72%, 1.72%, and 0% of all total creaky vowels. Having identified 88 parentheticals, Lee found that 75% of these were creaky. I identified 27 parentheticals across my four interviews, of which 13 (48%) were markedly creaky. Pearson’s Chi-squared test confirmed that there is statistical difference in the co-occurrence of creak with parentheticals between my data and hers: $\chi^2 = 5.7349$, df = 1, p-value = 0.0166. It would thus be worthwhile to revisit this issue again. It is possible, though, that Lady Gaga is an anomaly in that the majority of her parentheticals are generally not particularly creaky.

Additionally, the rather minor contribution of parentheticals to the total creaky voice vowels counts in my samples is also probably due to the fact that I identified very few
parentheticals in Lady Gaga's speech. I identified none in the Eur. sample, but it is unclear whether Lady Gaga naturally spoke with fewer parentheticals per minute than the consultants in Lee's study, as she does not indicate how long her interviews were, or if I was simply more conservative in the identification of parentheticals. As I noted in §5.3, I required parentheticals to have three of the defining characteristics to be marked as one. Lee (2015), while outlining these characteristics, does not specify how stringent her selection criteria was. She does note that peripheral content is to be expected in parentheticals. This may have been another trait she considered, which may have resulted in a greater number of parentheticals for her. There were a number of additional utterances in Lady Gaga's speech that, based on their content, seemed to be parentheticals, but did not exhibit three of the criteria and were thus not included.

The final structural factor considered was glottalization. The Ins. sample had the lowest percentage of creaky voice that appeared in glottalized segments despite have the highest percentage of total creaky vowels, while the Eur. sample had the highest percentage of creaky voice in glottalized segments, despite having the lowest percentage of creaky vowels ($\chi^2 = 72.68$, df = 3, p-value < 0.0001) (Figure 7).

In sum, the Ins. sample, which had the highest proportion of creaky vowels out of total vowels, had an equal proportion of creak appearing in the IP-final position and in
Figure 7: Percentage of creaky vowels appearing in a glottalized segment out of total creaky vowels

parentheticals as the other samples, except for the CFDA sample, which had the greatest percentage of creak appearing within parentheticals. The Ins. sample also had the lowest proportion of creak appearing within glottalized segments. Considered separately, it may be unclear what we can derive from these three bar graphs considering structural creaky voice in each sample. But, taken together, it becomes clear that Ins. sample has the highest proportion of non-structural creaky voice to structural creaky voice. The graphs were included to provide a visual, comparative breakdown of each structural factor in each sample. In the following section, I combine each of these factors to show how each sample differs in terms of total non-structural creaky voice to total structural creaky voice.

6.2 Structural and non-structural creaky voice across speech samples

I summed the number of creaky vowels in each sample that I identified as co-occurring with at least one of these three structural factors, and I found that the proportion of structural creak to non-structural creak is lowest in the Ins. sample ($\chi^2 = 47.28$, df = 3, p-value < 0.0001); it thus has the highest proportion of non-structural creaky voice. After considering these structural factors, then, we end up with the same distribution of creaky voice as we had in Figure 4, essentially: the Ins. sample has the highest proportion of non-structural
Figure 8: Percentage of non-structural creaky vowels out of total creaky vowels

creak, the CFDA and Nat. samples have relatively equal levels, and the Eur. pride has the lowest (Figure 8).

Table 3: Non-structural and structural creak across speech samples

<table>
<thead>
<tr>
<th></th>
<th>Ins.</th>
<th>CFDA</th>
<th>Nat.</th>
<th>Eur.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-structural creaky vowels</td>
<td>294</td>
<td>71</td>
<td>102</td>
<td>3</td>
</tr>
<tr>
<td>Structural creaky vowels</td>
<td>157</td>
<td>92</td>
<td>130</td>
<td>13</td>
</tr>
<tr>
<td>Total creaky vowels</td>
<td>451</td>
<td>163</td>
<td>232</td>
<td>16</td>
</tr>
<tr>
<td>Structural creaky vowels (%)</td>
<td>34.81</td>
<td>56.44</td>
<td>56.03</td>
<td>81.25</td>
</tr>
<tr>
<td>Non-structural vowels (%)</td>
<td>65.19</td>
<td>43.56</td>
<td>43.97</td>
<td>18.75</td>
</tr>
</tbody>
</table>
7 Meaning in Lady Gaga’s creaky voice

Given that we found speech sample to be a contributing factor to Lady Gaga’s differing usages of creaky voice, even after we had taken structural constraints into account, we might assume that Lady Gaga adopts creaky voice most extensively in the Ins. sample to realize some variety of socio-indexical meaning. To determine this meaning is no easy task, and it requires a movement beyond the quantitative. In §7.1, I consider Lady Gaga’s use of creaky voice within Labov’s classic Attention to Speech model (2001). I ultimately conclude that it cannot account for my data and is thus not the best framework for this analysis. In §7.2, I adopt a third-wave approach to stylistic variation, in which I claim that Lady Gaga employs creaky voice to take an affective stance of intimacy, and, further, to index character traits of openness and sincerity. I suggest that these traits may assist in the construction of a sincere pop-star persona.

7.1 Revisiting the Attention to Speech model

As noted briefly in §3, Labov’s Attention to Speech approach to intraspeaker stylistic variation prioritizes how aware speakers are of utterances they are producing. Unselfconscious speech is predicted to be more casual and vernacular, while self-conscious speech is expected to be more like the standard variety (Labov 2001). The goal of this approach is to determine speakers’ most natural way of speaking. Indeed, although this model has received criticism for its solitary emphasis on speech awareness and its misconception that speakers have only one authentic speech variety, it nonetheless has proved useful for Labov in accounting for data. Considering that Labov and his students have developed a variety of techniques to elicit both careful and casual speech and, given that Yuasa (2010) found in her perception study of creaky voice that it was perceived as “casual”, I would be remiss not to consider this model as a way to contextualize Lady Gaga’s creak usage across contexts.

One of the techniques used in sociolinguistic interviews to encourage interviewees to default to their “casual”, vernacular speech is to ease them into the divulging of a personal narrative (Labov 2001). As the Ins. sample was largely Lady Gaga relating personal childhood stories, one could argue that her increased use of creaky voice is simply reflective of her casual, unselfconscious style, and creaky voice is thus one of her vernacular variants. However, Lady Gaga tells stories of her youth and other personal anecdotes in her CFDA speech as well. In particular, relating the importance of fashion to her developing identity in high school, she talks about a designer coat she would long after from the window in a
consignment shop. She would check in with the shop owner every day to ensure he did not sell it to anyone, even though she knew she could never afford it. This story has no *markedly creaky* phrases. Note that I define markedly creaky segments as those that have at least three consecutive creaky vowels. If creaky voice were simply a vernacular, natural variant of her speech, one would expect it to occur extensively in this narrative.

In addition, that the CFDA sample (16% creak) and Eur. sample (1.5% creak) have such varying levels of creak is surprising, given that both are high-profile speeches. According to the Attention to Speech model, we would expect Lady Gaga to be paying attention to her speech in equal amounts. The CFDA sample is an acceptance speech in front of some of the country’s top fashion aficionados, and the Eur. sample is a gay rights speech at Europride. One could argue that Lady Gaga tells more narratives in the CFDA sample, resulting in the higher levels of creaky voice, although this is not necessarily the case. To determine the extent of narrative usage in each sample would be like splitting hairs, as both are structured around personal stories.

Ultimately, my results reveal that speakers can adopt different speech styles for contexts with equal levels of formality and casualness. Indeed, my claim here is not new; my data simply confirms what variationists (Eckert 2003; Estes-Schilling 2002) have suspected for some time. Due to these issues with the Labovian model, I adopt a third-wave approach to the analysis of intraspeaker stylistic variation in the following section.

### 7.2 A third-wave approach to meaning

As discussed in the literature review, intraspeaker variation in the third wave highlights speakers’ agency in conveying meaning through variants in speech. Speakers have more than one natural way of speaking which is not limited to the constraints of a casual or formal speech context. A speaker could choose to use vernacular features in a highly formal setting to achieve some social goal, for example. Determining these goals requires a consideration of the sociocultural landscape surrounding the variable at hand, as well the the speaker’s background and how they exploit it in interactional moments. While Labov’s (2001) model conceives of the social meaning of a variable as being either vernacular or standard, the third wave allows for more creativity on behalf of the speaker, thus requiring input from myriad sources to identify this meaning. Different approaches have been taken toward this end. Campbell-Kibler (2009; 2010) maintains that speaker perceptions are vitally important towards the determination of social meaning for variables. Podesva (2008) stresses the
importance of considering the discourse contexts in which variables appear. This study aims to combine both approaches. In lieu of my own perception study, I consider the results of Yuasa's (2010) survey. As for the discourse contexts in which creaky voice occurs, I draw on those from the Ins. sample, the sample in which Lady Gaga uses creaky voice most extensively.

Definitions of stance vary widely in the literature (Podesva 2013), but most consider it from a discourse perspective. Interpersonal stance is often defined as a person's relationship to their interlocutors in the exchange, and epistemic stance describes how certain the speaker is of her assertions (Theodoropoulou 2014). However, Podesva and Callier (2015) also note the possibility of considering the phonetic dimensions of stancetaking, particularly with regards to affective stance (the speaker's feelings toward a proposition), and how these might manifest through iconicity with physical manifestations of affect.

I find in the Ins. sample that Lady Gaga uses creaky voice most frequently when discussing emotionally-charged topics, such as her dead aunt, experiences with bullying, and her assiduous rise to fame. Indeed, the topic is always one that she is highly invested in. I thus argue that she employs creaky voice to take an affective stance of intimacy. Here, I define intimacy as an extreme closeness to or understanding of the topic of discussion. In other words, my claim is that she uses creaky voice when the topic at hand is highly personal to her in some way, and that she is affirming its intimacy to her through the adoption of creaky voice. Yet, as Podesva and Callier (2015) note, affective stances often become colored in context—they take on more specific meanings, which I will elaborate on later.

I consider 5 excerpts from the Ins. sample with markedly creaky segments, where I show why I think creaky voice indexes an affective stance of intimacy. Markedly creaky segments are those with three or more consecutive creaky vowels. As Podesva (2008) suggests, meaningful variation is not likely to be randomly scattered across swaths of utterances, but will cluster together, around significant contextual moments. Each numbered utterance in an excerpt represents an IP, and creaky syllables are bolded.

### 7.3 Creaky voice: Towards a stance of intimacy

In Excerpt 2, Lady Gaga discusses her late aunt Joanne after the interviewer happens to mention her. While Lady Gaga had not met Joanne before she passed, she is someone with whom she feels a deep connection, perhaps because of how much she meant to her father. There are three markedly creaky segments that are of particular interest to us: those in lines
6, 16, and 19-20. In line 6, we learn that Joanne was an artist, as Lady Gaga’s grandmother displayed her work around her house. In lines 19-20, Lady Gaga tells of how her aunt wrote poetry that never reached the point of publication. The link that joins these creak lines concerns art and its potential for visibility. Lady Gaga dropped out of New York University to pursue a path in music, but her talent was not appreciated immediately; indeed, she had been performing in hole-in-the-wall clubs and bars for about 6 years before she was signed by her first record label (Kristof 2012). She worked as a waitress to make ends meet, as her parents had cut her off. She truly understands the pain of unrecognized artistry and talent. In this way, then, she perhaps sympathizes with her aunt, whose artwork had only won the acclaim of her grandmother. In lines 19-20, Lady Gaga seeks to correct this deficit of fortune: she publishes her aunt’s poetry in her first album. They both are thus launched into the spotlight together.

I argue that in lines 6, 16, and 19-20, Lady Gaga reveals her intimacy with her aunt and with anonymous talent. Lady Gaga and her aunt are, in a sense, kindred spirits.

Excerpt 2: Aunt Joanne and her poetry

1. she was this enigmatic figure
2. in my life.
3. this person that died that
4. meant so much to my dad’s family.
5. I just felt really connected to her.
6. and my grandmother had her works of art all over the house.
7. I think that she in a way
8. makes us closer to our dad,
9. makes us understand him more.
10. because he’s very strict
11. and he’s very protective
12. and I think that losing her
13. contributed to that
14. ‘cause he just never wanted to let us go.
15. I printed um.
16. her poetry in my first album.
17. I remember I brought it to my dad
18. and I said she never got to be
19. a published author so
20. she is now.

In Excerpt 3, Lady Gaga’s and Joanne’s bond is further illuminated. Lady Gaga notes that she thinks her aunt is perhaps manifest in her (line 2) – an unsurprising revelation, given what we gleaned from the previous excerpt. As she divulges their spiritual connection, she notes how she and her crew engage in a type of ritualistic prayer before each performance, with her aunt at the center. Note the markedly creaky voice in line 5, ‘we say Joanne before we go on stage’. We might even assume that Lady Gaga invokes her aunt before her shows as she herself was an artist, and the shows represent a real-time embodiment of Lady Gaga’s performative talents and creativity. Of course, this is speculative, and the motivations for her invocation are irrelevant to the discourse context at hand. The significance of lines 5-6 is their revelation of Lady Gaga’s closeness to her aunt, and thus this topic of discussion.

**Excerpt 3: Aunt Joanne in prayer**

1. yeah I think she um
2. I think she lives through me.
3. every night
4. before the show when we pray
5. we say Joanne before we go on stage.

The conversation has progressed beyond the topic of her aunt in Excerpt 4, as Lady Gaga references her modest past growing up in New York City. Although not included in this excerpt, she notes elsewhere that her parents were the first in their families to go to college and were the children of Italian immigrants. She was not spoiled nor were her parents well-off in her adolescent years. She expresses how much her professional success means to her immediate family.
In this excerpt, we are offered a glimpse at how her professional success has improved the life of her immediate family. In line 2, she tells how her father at one time was not the most accepting man. The talk of her fans in line 4 leads me to believe she is referencing his rejection of homosexual lifestyles, as Lady Gaga is known for her popularity in the LGBTQ+ community. She notes in lines 5-7 that her sister’s and grandparent’s lives have changed as well, presumably for the better. In line 13, she reveals a personal aspiration: ‘to continue the legacy of her family’. Her true devotion to them, and their closeness and cohesion as a group – ‘I feel now like we just have this greater purpose as a family’ – thus manifests, elucidating the value of their ‘legacy’ to her.

Excerpt 4: Making her family proud

1. ya know my father wasn’t always like
2. the most open-minded person.
3. and watching his life change
4. meeting the fans,
5. watching my mother’s life change,
6. my sister,
7. my grandparents,
8. I feel now like we just have this greater purpose,
9. as a family.
10. at the heart of all this,
11. there’s someone very young and Italian,
12. deep inside me
13. that just really wants to continue the legacy of her family.

In Excerpt 5, Lady Gaga moves beyond family issues and describes the drop from her first record label. Laying on a couch, crying, she notes how emotionally distressed and dejected she felt. She notes her ultimate objective at the time – to be like Beyoncé. In line 2, Lady Gaga employs creaky voice in reference to Beyoncé just as she did when recalling Joanne: to proclaim their profound bond. With Joanne, this bond was linked to mutual feelings of untold talent. With Beyoncé, the connection is one shared goals. Creaky voice
occurs in ‘she’s a star’ in line 2 and ‘I wanna be on MTV’ in line 4, highlighting these two goals: to become a star and appear on MTV. For someone who had labored tirelessly to be signed by a record label only to be finally rejected, her unending ambition is a deeply-rooted one. Indeed, pursuing a music career was not just a side interest for her – it was real and constant, and one she would never let go. That she takes an affective stance of intimacy towards her aspirations thus becomes clear in lines 2 and 4.

Excerpt 5: Beyonce

1. I remember watching Beyonce
2. and thinking oh she’s, she’s a star.
3. I want that.
4. I wanna be on MTV.

In Excerpt 6, Lady Gaga continues to relate the difficulties she has faced. She elaborates on the theme of her album Born This Way in lines 1-5, which is that people can re-invent themselves and be re-born over and over until they have discovered their true identity. She draws upon her go-to assertion in line 5 – that each of us has the right to be who we desire – and overlays the clause ‘who you are’ with creaky voice in line 5. For Lady Gaga, and for her fans whom have experienced many of the same struggles she has, personal identity is almost sacrosanct; in this way, the stances of intimacy become evident. She does not take lightly nor for granted the fact that she now has the ability to perceive herself as beautiful, even if her features may not fit normative beauty conventions. In line 6, she confirms that she is, and always will be, an ‘underdog’. This inherent trait is not one that can be taken away from her – it is a part of her personhood.

Her conception of herself as an underdog has perhaps grown out of her troubled past. She tells of a traumatizing incident of bullying in high school in lines 7-13: a group of boys physically threw her in the trash, although we do not learn exactly why. Notice that creaky voice spans the entire utterance in line 10 – the statement actually addressing the event, as well as parts of line 6, where she references ‘the other’ girls, and all of line 7, where she notes that the locale of the incident was perhaps intentional, so that others ‘could see [her] in the trash’. This is the first excerpt of Lady Gaga using creaky voice when discussing an overtly negative situation. Zimman (2015) found that his speaker, a transgender male, adopted creaky voice extensively when distancing himself from highly
personal and traumatic experiences. One might argue, then, that Lady Gaga is using creaky voice heavily in this situation to achieve the same goal – to create a sort of barrier between her and this story, and, indeed the other girls (line 12) who authored it. Based on the surrounding discourse and the previous discussions, it seems unlikely that this would be the case, though. This interview serves as an inside look into Lady Gaga’s personal stories, hence the name ‘Inside the Outside’. Given that she also uses creaky voice to discuss issues which she is undoubtedly intimate with, such as her aunt, a stance of intimacy is just as likely here. I argue that she is embracing her past through an intimate stance, perhaps to convey a sense of sincerity and to align herself with her fans; I elaborate on this claim in §7.5.

Excerpt 6: Bullying

1. I define my own fame.
2. I define my own existence.
3. I define my own identity.
4. I define my own beauty for myself.
5. No one can define who you are.
6. I’m a perpetual underdog.
7. I went to meet some friends for some pizza
8. that were at the same pizzeria.
9. and the boys picked me up and
10. threw me in the trash can
11. on the street.
12. on the corner of my block while all the other girls were leaving
13. and could see me in the trash.

7.4 Intimacy and iconicity

As we have seen, Lady Gaga adopts creaky voice extensively for different functions – for example, to recall past events, divulge personal matters of importance, and offer metacommentary on her identity, but their overarching similarities involve her affective stance of
intimacy; she reveals her deep personal connection to the issues she is discussing, and it is creak that assists her in realizing this stance.

Indeed, I am not the first to suggest a meaning of intimacy for creaky voice (Ni Chasaide et al. 2004; Pennock-Speck 2005). Pennock-Speck (2005) contends that American female actresses employ creaky voice to increase desirability and intimacy; while his intended interpretation of intimacy refers to that of the sexual sort and mine of the emotional, the link is nonetheless clear.

One might wonder why creaky voice specifically is used in an intimacy stance. Previous literature has argued for phonation’s iconicity, or rhematicity, with specific physical manifestations of affect (Podesva and Callier 2015). Tense voice has been associated with stress, anger, and confidence, for example, and low harmonic energy is often attributed to sadness or tenderness. As Zimman (2015) notes, creaky voice restricts the amplitude and fundamental frequency range to which speakers have access. I argue that this vocal constraint relays a quality of rawness and bareness, one devoid of high pitch, and perhaps intonational, fluctuations. As affective intimacy is characterized by a reserved and minimized affect, and thereby rawness and bareness for its lack thereof, it is reasonable that she would use creaky voice to convey this meaning.

Recent work (Zimman 2015; Levon 2015) has noted that creaky voice might serve as a verbal tool for speakers to distance themselves from certain propositions in speech. Specifically, Zimman notes that his consultant uses creaky voice to distance himself from topics which are of a highly personal nature. Levon (2015) claims that his speaker’s usage of creaky voice is reflective of an emotional tension, as he endeavors to reconcile his sexuality within the teachings Orthodox Judaism. While their interpretation on a surface level appears to challenge my sentiment of creaky voice’s link to affective intimacy, there are indeed similarities between our analyses. In both cases, speakers seem to have a tendency to use creaky voice alongside private issues. Whether a speaker is expressing distance or intimacy is up to the interpretation of the ethnographer.

7.5 Openness and sincerity

Podesva and Callier (2015) argue that affective stance and variable iconicity are important beginning steps to consider in discerning the social meaning of non-modal phonation. As Kiesling (2009) argues, “Identity and personal style are both ways of stereotyping habitual patterns of stancetaking, or repertoires of stance.” Kiesling’s claim suggests, then, that the
repeated use of a particular stance can assist in the construction of a style. Given that I have identified 5 different instances of Lady Gaga using creaky voice to take an affective stance of intimacy in the Ins. sample, I argue that the continued reappearance of this stance is salient enough to compose a style, or persona – that is, one of a sincere pop star persona. In the following paragraphs, I draw on Lady Gaga’s personal and professional background, as well as theories of fandom and persona construction, as I seek to prove this claim.

A particular affective stance – such as intimacy – may index an attribute, or fleeting personality trait. For example, if someone were to express uncertainty about a certain proposition, this stance might index a personality trait of powerlessness (Jaffe 2009), meaning listeners would interpret the speaker as possessing this trait. I contend that Lady Gaga’s usage of creaky to take a stance of affective intimacy indexes the openness and sincerity of someone willing to share the details of personal struggles. As I noted, creaky voice has a raw vocal quality devoid of extreme fluctuations in pitch and intonation. This bareness is reflective of her desire to bare all – to be open, devoid of a facade and guarded veneer. As she divulges these details, I argue, she is aiming to align herself with her fans, many of whom share similar experiences. Indeed, Dilling-Hansen (2015) found in her case study of Lady Gaga’s fandom that in Lady Gaga, fans find a kindred spirit which bonds them. She has quoted a fan as saying, “When you see her and she sings about it [personal struggles] and talks about it, she’s sharing it with millions of people. And she knows we all go through it. And you start crying because you realize things can get better.”

As we have seen, her propensity to share personal struggles was particularly salient during the Ins. sample, which was filmed several months after the release of her third studio album – Born This Way. As the name suggests, this album preached self-love and the called for social equality amongst marginalized groups, particularly the LGBTQ+ community and those having dealt with or currently dealing with harassment and mental health issues. The purpose of this interview was to offer an inside look at her troubled past and how it has influenced her identity as a pop star and the development of her new album. When asked in an interview filmed less than a month prior to the filming of the Ins. sample if she uses her personality to sell her music, she remarked, “It’s kind of like asking, ‘Are you honest?’ Do you have integrity? Are you full of shit?’ But listen, my social involvement for equality and gay rights, as well as my music, is completely genuine. I have no reason to do or say anything unless it comes from my soul” (Vena 2011). Her focus on aligning herself with her fans – many of whom were a part of LGBTQ+ circles even before her outspoken love for the community and initiatives addressing social inequality – thus spawned during this
time. Whether the stories she tells and allegiances she professes are actually sincere or not is irrelevant. Her persona of a real, open, and genuine pop star is one she fashions intentionally for a specific social end: to garner the trust and devotion of die-hard fans and other casual listeners. She utilizes the rawness of creaky voice to embellish the rawness associated with an undisguised personal essence. Dilling-Hansen (2015) explicitly notes that Lady Gaga’s sincerity and genuineness are vital to the healthy relationship between her and her fan base. In fact, as it seems from Dilling-Hansen 2015 that Lady Gaga is steadfast in maintaining this quality of sincerity, it might be part of a more concrete identity rather than a fleeting persona, which she chooses to build upon most in the Ins. sample.

It is worth noting that the persona Lady Gaga constructs is perhaps a foil to those of other pop stars – “by definition, pop stars are manufactured”, claims New York Magazine (Grigoriadis 2011). Indeed, the stereotypical female blonde pop star lacks substance and is contrived, fabricated, and marketed by the music industry solely for profit. She is not sincere or genuine – her voice, music, and image are altered until culturally specific perceptions of perfection are virtually attained. While retaining the expected beauty norms of celebrities, Lady Gaga nonetheless seeks to differentiate herself from these other pop stars and celebrities by portraying herself as someone just as haunted as everyone else, just as grounded, and just as passionate. Considering that participants in Yuasa’s (2010) perception study rated creaky voice as sounding more genuine than modal voice, it is unsurprising that she uses it to realize these identity-based meanings.

I would like to conclude this discussion by referencing one possible pitfall of my analysis: personae construction has often been found to be mediated through a combination of multiple variables in speech, not just one. Podesva (2008) notes that linguistic styles consist of a variety of different linguistic features, the social meanings of which are synthesized in a style. While I have indeed found that creaky voice usage varies across different speech samples, suggesting that Lady Gaga employs different styles in each sample, a consideration of other features in addition to creaky voice could lead us to a different stylistic conclusion. For example, Campbell-Kibler (2011) has shown that one variable can influence our perception of another. She notes, “-/s/-backing in the case of an ambiguously Southern speaker promotes a perception of him as ‘country’ and shifts the relationship between competence, sexuality and masculinity, by increasing perceptions of the complex style ‘masculine, unintelligent, straight man’” (2011:64). Of course, my study is not the first to argue that a sole variant might assist in the production of a persona (Podesva 2007), and, the repeated use of a variable co-occurring with a particular stance can contribute to a certain speech style.
8 Beyond Lady Gaga: Indexical field for creaky voice

As noted previously, a variety of social meanings have been proposed for creaky voice. While Podesva (2013) argues that meanings of variables will differ substantially across different discourse contexts and populations, there is sometimes an underlying meaning from which all others are ideologically linked (Eckert 2008). This system of meanings, which also was referenced earlier, is known as an indexical field.

Podesva’s (2013) indexical field for creaky voice places at its center the phonation’s similarity to men’s voices. He argues that this meaning can explain Yuasa’s (2010) interpretation of creaky voice as indexing a feminine “upwardly-mobile persona” which is capable of competing with male peers, as well as Mendoza-Denton’s (2011) claim that for a group of gang-affiliated Latinas, it is used to portray a hardcore persona, and Leftkowitz and Siccoli’s (2007) assertion that the group of women in their study used creaky voice to convey authoritativeness. Although his analysis works for the meanings he discusses, it is rather unclear how a “similarity to men’s voices” might account for perceptions of creaky voice as sounding intimate (Pennock-Speck 2005; Ni Chasaide et al. 2004), relaxed, or bored (Ni Chasaide et al. 2004). Similarly, how might Zimman’s (2015) interpretation of creaky voice as a distancing tool arise out this proposal?

After considering Lady Gaga’s use of creaky voice in addition to each of these proposed meanings, I adopt Levon’s (2015) claim that creaky voice broadly indexes “contained” emotion. First, this analysis can account for my data. Lady Gaga might employ this suppressed-affect meaning of creaky voice to take a stance of intimacy as the notion of intimacy itself, in a sense, is characterized by seriousness and a lack of emotional embellishment—an investment in something so deep that extreme emotional displays accompanying it might be interpreted as contrivances. Additionally, it can quite clearly account for Zimman’s (2015) claim that creaky voice can be adopted to detach the speaker from some person or object, as emotional distance is characterized by a lack of emotional engagement as well. Even further, this interpretation can explain the proposed meanings of authoritativeness, toughness, and professionalism. These three character traits are often associated with men, and men are stereotypically deemed the rational sex, not driven by emotional impulses. For these reasons, then, women may seek to convey toughness, professionalism, and authoritativeness through creak’s constrained-affect meaning. Finally, it is clear how this notion of sustained affect might index boredom and relaxation (Ni Chasaide et al. 2004), an affective stances arising from minimized emotional engagement.
with outside stimuli. Under this overall interpretation, creaky voice’s core meaning is linked to its iconic potential, and its other meanings arise from ideological associations with this iconicity.

9 Conclusion

In this thesis, I have adopted a third-wave approach to an intraspeaker analysis of creaky voice. Having considered the speech of Lady Gaga across four different speech samples and in the discourse contexts within them, I have argued that in the “Inside the Outside” interview, Lady Gaga employs creaky voice to signal affective intimacy to the topics at hand. Further, I have argued that stances of intimacy can index character traits of openness and sincerity. Ethnographic work into Lady Gaga’s relationship with her fans has revealed that she may be using this openness to align herself with them and their struggles, as she constructs a sincere pop star persona-based identity – that of a pop star who seeks to bond deeply with her fandom through shared personal struggles. While this persona seems to manifest most prevalently in the “Inside the Outside” interview, it is nonetheless one that distinguishes her from other pop stars more generally. We might contend, then, that this contextual persona is one that contributes to a more static identity of the same sort. As for the meaning of creaky voice in a wider socio-cultural landscape, I have adopted Levon’s (2015) interpretation of its sustained-affect meaning. I argue that this meaning can explain other’s proposals through ideological connections embedded in the indexical field.

This thesis as adding to the growing body of research combining quantitative, discourse analytic, and linguistic anthropological methods to the study of sociolinguistic variation. It highlights the importance of adopting ethnographic approaches as we examine stylistic practice and social meaning, as the more profoundly we understand how a proposed variable situates itself in a speaker’s life story, the better we can discern its stylistic purpose, not only for the speaker, but for others, as well. While I am not claiming that all meanings composing an indexical field are activated for every speaker in every interaction, variation beyond the level of phonology may draw core meanings from iconic manifestations, as is true with creaky voice, and thus there will undoubtedly be similarities. Sicoli (2015) dismisses the importance of considering broad, “abstract social meanings” of variables, noting that it would be more worthwhile to situate meanings with smaller socio-cultural landscapes. However, his claim fails to acknowledge the undeniable iconic meanings that speakers may associate with variables, which circulate beyond situational social relations.
This notion is superbly reflected in the indexical field for creaky voice, which has a vast network of meanings that can ultimately be traced back to one source – its sustained affect meaning, rooted in creaky voice’s iconicity with certain affective stances.

Before offering some final thoughts, the female celebrities referenced in the beginning of this study deserve some recognition, as well: Kesha, Britney Spears, and Kim Kardashian, and Katy Perry. Future work might consider how these women use creaky voice stylistically. Of course, we would be naive to believe that they employ creaky voice to achieve the same goals as Lady Gaga. Clearly, as I have shown, even Lady Gaga differs in her use of creaky voice across contexts. I suspect we would again find that their usage is rooted in the sustained-affect meaning, though. In a short YouTube video on creaky voice and the Kardashians’ use of it, a woman noted that they try to sound “cool” and like they “don’t care about anything” (MightFalcony2011 2013). Indeed, the character type of the cool kid with a measured demeanor, void of enthusiasm, is a familiar one, and it is evident how this “coolness” might be channeled through creak’s sustained-affect meaning.

Overall, this study offers a first foray into the sociolinguistic examination of the speech styles of American female pop stars. On the most basic level, it is clear that many pop stars undergo regular aesthetic transformations to convey something new to the public. Lady Gaga, for example, was once infamous for her outrageous fashion choices. Recently, she has been dressing much more mature, perhaps because she had been receiving a lot of criticism for solely trying to “shock”, or perhaps because she has been recording jazz records with Tony Bennett and seeks respect from his undoubtedly older fan base. Katy Perry’s transformation over the years has moved in almost the opposite trajectory. I argue that their speech styles serve as a sort of verbal aesthetics choice, serving to reflect, reinforce, and create these new identities and personas. Through studies like this one, we can learn how celebrities situate themselves within the larger pop cultural sphere; more broadly, we can learn how non-celebrities mark a place in their localized stylistic landscapes, as well.
Appendix A

Inside the Outside interview
https://www.youtube.com/watch?v=Gjt-EW3QGuK

The National interview
https://www.youtube.com/watch?v=6ouFWXdpnpI

CFDA acceptance speech
https://www.youtube.com/watch?v=HLWQMZYEZqk

Europride speech
https://www.youtube.com/watch?v=HyB9qsRNjIM

MP4 to MP3 converter
http://www.clipconverter.cc/
References


