LIFESPANS & STYLES
Undergraduate Papers in Sociolinguistics

Volume 6 Issue 2

2020

Podcaster Prosody: Creaky Voice and Sarah Koenig's Journalistic Persona

Edgar Yau
eyau3@u.rochester.edu

ISSN: 2057-1720 doi: 10.2218/ls.v6i2.2020.5215

This paper is available at: http://journals.ed.ac.uk/lifespansstyles
Hosted by The University of Edinburgh Journal Hosting Service: http://journals.ed.ac.uk/

Podcaster Prosody: Creaky Voice and Sarah Koenig's Journalistic Persona

Edgar Yau

This paper addresses the style-shifting of podcast host Sarah Koenig, specifically in her use of utterance final creaky voice in different contexts. I find that Koenig uses more creaky voice on her podcast *Serial* than in an interview context. Additionally, her creaky voice in the interview occurs in specific contexts related to her work as a journalist. Based on analyses of how phonetic features can construct certain personae, I argue that Koenig may be designing her speech to construct a journalistic persona with her use of creaky voice.

1 Introduction

Podcasts are episodic audio series that are typically downloaded and streamed. In purely auditory mediums like podcasts and radio, hosts must connect with their audience through only their voice. While there are many linguistic studies about radio and radio hosts (e.g., Rodero 2014), there are very few studies on the language of podcasting and podcast hosts. In 2014, the creators of the popular American radio show *This American Life* uploaded the first episode of the podcast *Serial*, created and hosted by journalist and radio personality Sarah Koenig. Years after its initial release, *Serial* holds the record for the most listened to podcast of all time and had been downloaded over 320 million times as of its third season in late 2018 (Quah 2018). The popularity of the show and its format have since been widely credited for popularising podcasting as a medium, sparking what has been referred to as the "Great Podcast Renaissance" (Roose 2014). Koenig's voice has a distinctive prosodic quality, spurring many performers to mimic and parody her delivery style, from the mockumentary *American Vanda* to the comedy sketch show *Saturday Night Live*. A public figure with a distinctive speech style like Sarah Koenig invites linguistic study. One recurring feature in every parody of Koenig or piece of commentary on her speech is her prolific use of utterance final creaky voice.

Creaky voice, sometimes referred to as vocal fry or glottal fry, is a distinctive vocal quality characterised by the use of the lowest part of one's vocal register. Most often, creaky voice occurs at the end of utterances, but it can occur in other positions in the utterance as well (Wolk et al. 2012). It is a widely observed, commented on, and investigated feature, with many available descriptions of the phonetic features of speakers with creaky voice. With this study on intraspeaker variation, I hope to show an example of how creaky voice can be used by speakers to construct their personae. Studies into the stylistic use of prosodic features like Podesva (2007) provide a basis for my account of Koenig's use of creaky voice to construct a social identity. In this paper, I analyse the social meaning of creaky voice and how Koenig's use varies across contexts.

My hypothesis, based on past studies on intraspeaker variation, is that Koenig style-shifts between her interview speech and podcasting speech. I predict that she increases her use of creaky voice in the podcast context to index authority and construct a journalistic persona. The qualities of this journalistic persona are based on a study by Gasser et al. (2019), which looks into the production and social meaning of distinctive features of newscaster prosody in American English. I will discuss this and other studies more below. Ultimately, I find that her use of creaky voice in the interview context lines up with instances where she is making emphatic and authoritative statements about the role of journalists and her job.

1.1 Creaky Voice

In this case study, I use perceptual studies of creaky voice as the basis for my understanding of the social meaning of the feature in contemporary US English. Wolk et al. (2012), Ligon et al. (2018), Yuasa (2010), and Parker and Borrie (2017) offer differing accounts of perceptions of speakers using creaky voice.

Wolk et al. (2012) recorded 34 female native American English speakers between the ages of 18 and 25. The participants were asked to produce three sustained vowels and then read passages of text aloud. Two speech pathologists were then asked to listen to the recordings and judge whether tokens were produced with or without creaky voice and where in the sentence they perceived the creaky voice to occur. Wolk et al. (2012:e114) found that: "(1) vocal fry was used in sentence reading by more than two-thirds of this population of female college students, (2) vocal fry rarely occurred in sustained vowels, (3) vocal fry occurred most often at the end of utterances, and (4) statistically significant differences were found for several acoustic measures between vocal fry and normal register". The key takeaway from this study is that creaky voice occurs most often in the utterance final position, which seems to be reflected in Koenig's speech.

As a feature, creaky voice is heavily associated with women. Studies have shown that women, particularly young women, use creaky voice more than men, and it is far more commented on and criticised in women. Higdon

(2016:212) provides examples of public perception of creaky voice being "overwhelmingly negative", citing an open letter written in *The Guardian* titled "Young women, give up the vocal fry and reclaim your strong female voice" (Wolf 2015). In the article, Wolf (2015) describes vocal fry (creaky voice) as "that guttural growl at the back of the throat, as a Valley girl might sound if she had been shouting herself hoarse at a rave all night". It would seem, then, that creaky voice is not considered a "positive" feature when noticed in women's speech.

Ligon et al. (2018:805.e25) gave 23 participants a set of adjectives to mark on a scale of desirability, on the basis of which they classified a set of adjectives as "undesirable", including "apathetic/disinterested, unhealthy/sick, bored/unengaged, sleepy, vain, depressed, and obnoxious". They then played the participants recordings of eight female speakers with different "vocal quality profiles", one of which was vocal fry. They found that 32% of the participants used only "undesirable" adjectives to describe the speaker with vocal fry, whereas just 4% used only "desirable" adjectives; the rest of the 64% used a mixture of "desirable" and "undesirable" adjectives to describe the speakers. While more participants used only negative adjectives to describe the speaker with vocal fry than those who used only positive, the majority used a mixed set of adjectives.

Further studies have shown that the perception of vocal fry can be affected by other prosodic features. Parker and Borrie (2017) conducted a study on listeners' judgements of female speakers with and without vocal fry. By asking participants to rate speech samples from eight English-speaking women differing in pitch, speed, and the presence or absence of vocal fry, they found that in lower-pitched speakers (Average F0 around 172.49 Hz) who spoke quickly (5.39± 0.24 syllables per second), the presence of vocal fry decreased the "likeability" and perceived intelligence of the speaker. However, in lower-pitched speakers who spoke slowly (3.84± 0.09 syllables per second), speakers with vocal fry were rated marginally more "likeable" and more intelligent than speakers without it.

In another perception study, Yuasa (2010:328) played participants two US English speech samples from the same female speaker, one speech sample containing a "considerable amount" of creaky voice and the other with mostly noncreaky voice. The 175 college student participants were first asked to identify whether they thought women in their area used considerable creaky voice. The participants were then given sets of opposing adjectives (e.g., knowledgeable/ignorant) and asked to rate the speech sample on a scale of one or the other. Yuasa (2010) found that 78.9% of participants said they heard creaky voice often used by women in their area indicating that creaky voice is a widespread phenomenon for American women. She also found that participants were likely to perceive creaky voice as "educated", "genuine", "casual", and "compliant", and that there was no real difference in perceived confidence. These traits, as opposed to "ignorant", "pretentious", "formal", and "aggressive", respectively, have generally positive connotations. Ward (2006:151–153), who Yuasa cites, also finds that creaky voice occurs when utterances are "intended as authoritative statements, advice, opinions, decisions, recollections, etc.". From her study, we can see that creaky voice can have a socially mobile, educated social meaning for American women.

These accounts of the perception of creaky voice provide a range of conclusions about the potential social meaning of the feature. While some studies and commentaries identify it as a vocal quality with negative effects on perception of the speaker, the majority describe the results as mixed. Overall, perceptions of creaky voice can range from apathetic, to lascivious, to upwardly mobile, and to authoritative. In this study, I will focus on my intuition of how a female journalist might use creaky voice to access this authoritative social meaning.

1.2 Style-shifting and Constructing Personae

As the basis of my analysis of Koenig's style-shifting, I draw on sociolinguistic studies that focus on the use of linguistic features to construct specific personae and identities. Following Coupland's (2002) theory of style and identity, sociolinguists have done a variety of studies on the social meaning of language and how that meaning intersects with the identities of speakers. Coupland (2002:186) proposes a theory of style that "can explore the role of style in projecting speakers' often-complex identities and in defining social relationships and other configurations of context". I use this model to analyse Koenig's use of different linguistic features as creating different identities. Previous studies such as Eckert (2011), Podesva (2007), and Zhang (2005) provide a basis for this type of study and include a variety of examples of how linguistic variables can be used by speakers to index certain features and construct different personae.

Eckert (2011) describes the style-shifting of a preadolescent girl, Rachel. She found that a fronted nucleus in the /ow/ vowel was associated with "trendiness" and "adolescence". Rachel's /ow/ vowel changed depending on context and subject matter. Eckert concluded that the preadolescents were able to draw on the connotations of sound changes and would exhibit the sound change to index different identities. In Rachel's case, she exhibited a fronted /ow/ to index a teenage stance that might set her apart from other kinds of girls and female personae. In her analysis of Rachel, Eckert (2011:85) frames Rachel's style-shifting within the context of the linguistic marketplace, specifically the "heterosexual market". This framework uses social capital as one explanation for style-shifting, e.g., one might style-shift into a different linguistic register to gain social standing, whether that be using the features of an in-group or other features seen as desirable under certain circumstances. Zhang (2005),

for instance, describes the use of phonetic features in the context of social change and class stratification in Beijing. In her study, she analyses the speech of 14 upwardly mobile Chinese "yuppies" and 14 state professionals. She looks at phonological features that index three specific Beijing Identities: the "alley saunter", the "smooth operator", and the "cosmopolitan". She concludes that young professionals in Beijing use linguistic resources to identify themselves as cosmopolitan to stand out in the job marketplace. Additionally, she characterises the variation occurring in Mandarin Chinese in this youth demographic as the utilisation of linguistic features in different contexts for different purposes of indexicality and identity construction, rather than as a matter of standard versus non-standard language.

Podesva (2007) looks specifically at how intonational variation can be used to construct personae. The paper investigates how a single speaker, Heath, uses falsetto across different contexts. Podesva (2007:496) found that Heath exhibited "significant variation across situations" in his use of falsetto phonation. At a barbecue with close friends, Heath's falsetto was "more frequent, longer in duration, higher in f0, and wider in f0 range" than in a conversation with his father and in a work context (Podesva 2007:497). Additionally, by looking closer at the particulars of the conversations in which falsetto occurred, Podesva (2007:497) found that Heath was likely using falsetto to "construct a diva persona and perhaps a gay identity". In his case study, Podesva makes a compelling case for the use of voice quality as a stylistic variable, how semantic context can affect linguistic variables, and how that semantic context may be analysed to infer what identity might be constructed by the variable in question. He identifies creaky voice as a stylistic feature of Heath's speech used to widen his pitch range in the construction of this diva identity. The work of Podesva (2007), Zhang (2005), Eckert (2011), and others demonstrates that people are both aware and adept at using linguistic features to construct personae. Additionally, we can draw from these articles that case studies are a useful way of examining style-shifting. Thus, a case study of Sarah Koenig's style-shifting may reveal new insights into stylistic trends.

1.3 The Speaker, Sarah Koenig

Sarah Koenig is a journalist, radio personality, and podcast host from New York City. After a career as a print journalist, Koenig began work in radio broadcasting, working as a producer and occasional host on popular radio show *This American Life* with frequent collaborator Ira Glass. Following the release of *Serial* in 2014, Koenig won numerous awards for the show, including the first Peabody Award ever to be given to a podcast (Barclay 2020).

The first season of *Serial* follows Koenig as she attempts to uncover evidence in a controversial murder trial. Following the success of the first season, the show creators then went on to make a second season following the case of Bowe Bergdahl, a U.S. Army soldier accused of being a deserter after being held captive by the Taliban for 5 years. Koenig was 45 years old at the time of recording the second season.

In a blog post, Mark Liberman (2015) looked at two speech samples of Koenig, one from 2000 and one from 2014. He plotted her distribution of f0 values and found that "in 2014, Koenig is using the lower part of her range (100–150 Hz) more extensively; and is also dropping into a period-doubling register (40–80 Hz) to a much greater extent". Liberman (2015) proposes that "Sarah Koenig's vocal fry seems to be something new" and suggests that she is following "a general stylistic trend". He further speculates that the change in her style may have been influenced by close collaborator Ira Glass, with whom she worked on the radio show *This American Life*. While Liberman's quick look is limited, it is evidence that there has been a change over time that is more than just the natural lowering of pitch that comes with age. Here, I argue that creaky voice is a feature that Koenig uses agentively across different contexts. I chose Koenig as the subject because of the aforementioned observations on and criticisms of Koenig's use of creaky voice. Luckily, there is no shortage of Sarah Koenig speech samples to study.

2 The Journalistic Persona

In the past, Koenig has said that she uses "different voices" in different contexts. In an interview, Koenig responds to a question about a difference in her voice in the first and second seasons of *Serial* (Point Park University 2016). She identifies her Season 2 voice as more "traditional journalist" and describes how using a more "girl sleuth" voice, as she did in Season 1, would have felt "disingenuous". Although she describes the change in her voice as not a conscious decision, she is aware of her different vocal personae. To understand what personae Koenig may be constructing or accessing, it is helpful to understand her goals in speaking.

In her article "Finding a pitch that resonates: an examination of gender and vocal authority in podcasting", Mottram (2016) identifies how women index authority in the medium of podcasting. While not a phonetic study, Mottram (2016) identifies authority as a primary goal of podcasters in their speech. In a chapter about "indexical obsolescence", Eckert (2020) identifies "The *Marketplace* Register". This style, which she claims is used by broadcasters on NPR show *Marketplace* and elsewhere, uses a "good deal" of creaky voice (Eckert 2020:5), which

marks the style as more youthful. She describes the *Marketplace* register as showing "more commentary and a more immediate engagement with the audience and action" than traditional, and more conservative, newscaster speech (Eckert 2020:4). She also found that the style was "linked to economic stories that are serious, but with a light key, and even a trendy edge", which contrasts with the older, more traditional newscaster register (Eckert 2020:10). Ira Glass, host of radio program *This American Life,* is quoted in the chapter saying, "[i]f older people like their news stories delivered in a detached style, younger people prefer stories with some engagement" (Eckert 2020:15). Glass recognises an overall stylistic trend in radio broadcasting that is more youthful and engaging. As mentioned above, Glass and Koenig are close collaborators, so their understanding of these different broadcasting styles may be similar. Podcasting is a new medium geared towards a younger, more technologically savvy audience; thus, a register that uses creaky voice, which indexes both youth and authority, would be appropriate for the emerging medium.

Based on the intuition that newscasters have a distinctive and recognisable prosody, Gasser et al. (2019) looked at the prosody of newscaster speech, specifically the pitch contour of their declarative utterances. Gasser et al.'s (2019) study consists of a production and perception experiment. For the production experiment, they compared two sets of non-newscaster volunteers reading 12 sentences taken from previous news broadcasts. One set had the original script and knew that they were reading news broadcasts, and the other set had a modified script that placed the sentences in conversational contexts designed to minimise any intonational effect. The perceptual experiment had participants listen to recordings and identify whether or not these recordings were "newscaster" or "everyday" speech. Their studies found that newscasters use distinct prosodic features that listeners can identify and perceive as "newscaster-like". Additionally, they surveyed 12 newscasters and found that the qualities that newscasters want to index are "trustworthiness" and "engagement", as well as "likeability, intimacy, and authority". Thus, it would make sense that the linguistic identity of newscasters would be tied to certain features that index credibility, with a certain storytelling quality in pursuit of engagement. Because people recognise this style of speech as "newscaster-like", it is possible that newscasters and reporters might design their speech to inhabit this trustworthy newscaster persona.

The particular feature of newscaster prosody identified by Gasser et al. (2019:7) that I am concerned with is: "Newscasters will use wider pitch excursions and lower L% targets to convey authority". In the ToBI transcription system, L% describes a low boundary tone, i.e., a low pitch at the boundary of an utterance. As creaky voice "occurs typically when speakers lower their vocal pitch to the lowest register they are capable of producing" (Anderson et al. 2014), creaky voice would be a good indicator of the wider pitch excursions and lower targets that Gasser et al. (2019) describe. Additionally, the lower L% targets described coincide well with the finding in Wolk et al. (2012) that creaky voice most commonly occurs at the end of an utterance. According to these findings, we understand utterance final creaky voice as a feature that can be used to convey authority.

As an auditory communicative medium, podcasting is similar to radio. Rodero (2014) performed a study on intonation in radio news. She looked at 24 radio broadcasters in Spain and in analysing their pitch found that newscasters employ regular pitch contours and constant emphatic stress. While her perceptual study found that this intonation is often ineffective in engaging listeners, her findings still reveal that newscasters, regardless of language and medium, design their speech using prosodic features to engage listeners.

Serial falls under the genre of the "true crime" podcast, and Koenig's role in *Serial* is similar to a reporter or a newscaster. She wants to tell a story with a level of credibility and authority while engaging an audience. Understanding Sarah Koenig's history as a journalist and radio personality helps us understand what she might specifically be trying to index. Based on these analyses of radio personalities and newscasters, I identified that the qualities newscasters want to index are engagement and authority: a speaker might shift into a journalistic identity to index these traits. In this paper, I combine analyses of creaky voice with studies on the prosody and goals of news presenters as the basis for my analysis of how Sarah Koenig uses creaky voice. My hypothesis is that Koenig will use significantly more utterance final creaky voice in the podcast to index an authoritative journalistic identity.

3 Methodology

This study looks at declarative utterances over 10 minutes of speech in each context. I divided the speech samples into utterances by listening to the audio, then labelled each utterance "Creaky voice" or "No Creaky Voice". I identified creaky voice by listening to the recordings and verified my findings using the pitch tracker built into Praat (Boersma 2001). As described in Yuasa (2010:324), the spectrogram displays "vertical striations spaced well apart indicating low-frequency vibrations", which were often identifiable through irregular breaks in Praat's pitch tracking algorithm. The lowest end of Koenig's pitch range without creaky voice is 130Hz, however with creaky voice, she dips below 110Hz to as low as 80Hz, so any utterance that ended below 110Hz was counted. An utterance was not counted as an occurrence if my judgements conflicted with the spectrogram information. Figure 1 shows an example of an utterance with utterance final creaky voice. Compared with the "No Creaky Voice" utterance in Figure 2, Figure 1 shows an excursion into the lowest pitch range, greater vertical striation, and irregular breaks in the pitch tracking algorithm.

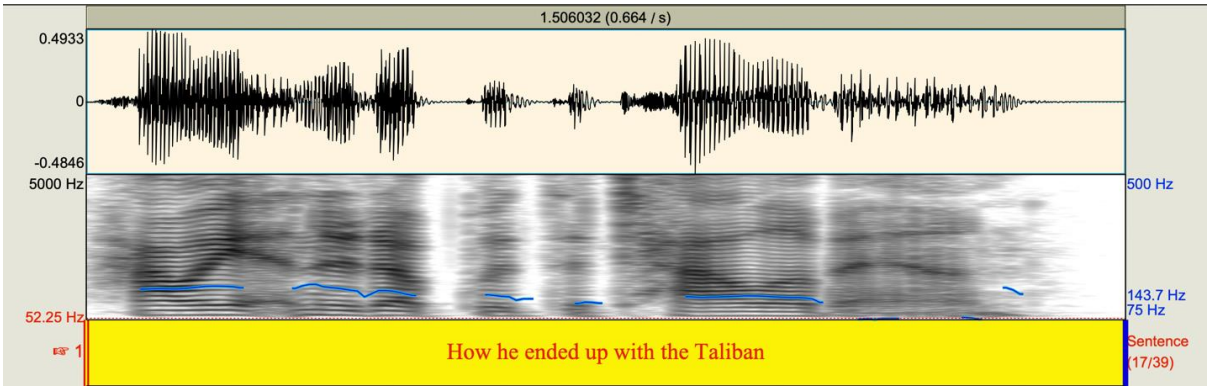

Figure 1: Example of utterance final creaky voice from *Serial* (Koenig 2015).

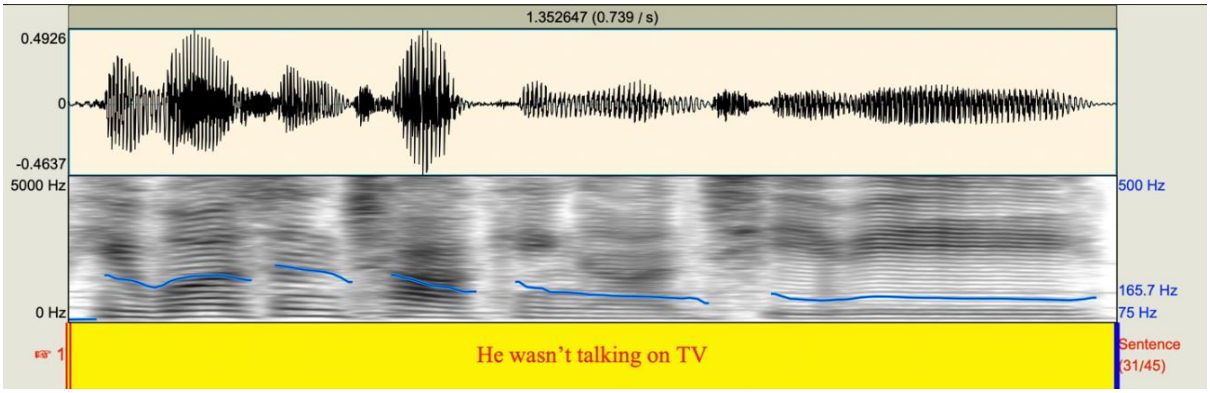

Figure 2: Example of an utterance without creaky voice from *Serial* (Koenig 2015).

The data was collected from two different sources: an interview done with Koenig in 2017, "The Fireside Interview" (Qualtrics XM 2017), about the second season of *Serial* and podcasting, and Season 2 Episode 1 of *Serial* "Dustwun" (Koenig 2015). These contexts were chosen because they were held within 2 years of each other and because the interview directly addresses the second season. For this study, only the initial 10 minutes of her speech were considered.

The content of the two contexts varies greatly. The episode of *Serial* is the introduction of the second season. In the episode, Koenig describes the basic facts of the case of Bowe Bergdahl, interspersed with news clips and interviews with Bergdahl himself (not conducted by her). One feature of note is that Koenig never actually has any contact with Bergdahl. As such, all the declarative utterances counted are Koenig speaking to the audience and are most likely prewritten.

In the Fireside Interview, Koenig is speaking with an interviewer in front of an audience about the making of *Serial*, her job as a journalist, the differences between the first and second seasons of *Serial*, and the emergence of the medium of podcasting in the mainstream. As Podesva's (2007) study of Heath showed, content and context may reveal the social meaning of the feature. So, after identifying utterances with and without creaky voice, I transcribed sentences that showed varied creaky voice distribution and noted whether the semantic content of each utterance was related to journalism or reporting and her experience in the field.

The primary uncontrolled factor between the two contexts is the presence of the audience and the interlocutor. While both *Serial* and the interview are directed towards an audience, the presence of a live audience might affect Koenig's speech. On the other hand, the presence of a "known" audience in both contexts may also provoke similarities between the contexts. The interlocutor is an additional uncontrolled factor, but his presence throughout the interview is minimal. He rarely interjects and only does so to ask questions. With these uncontrolled factors in mind, the data below shows differences in quantity between the two contexts.

4 Data

Table 1: Percentage of utterances with final creaky voice in different contexts

Context	% Creaky voice (*N*=206)	% No Creaky Voice (*N*=240)	Total Count
***Serial* S2E01**	64% (139)	36% (80)	219
Fireside Interview	30% (67)	70% (160)	227

The data in Table 1 includes all declarative utterances over the initial 10 minutes of speech in each context (20 minutes of speech in total). Over the 10 minutes of speech, a total of 446 declarative utterances, 219 from the *Serial* episode and 227 from the "Fireside Interview", were counted and labelled "Creaky Voice" or "No Creaky Voice" depending on the spectrogram information shown in Praat and my own judgements, as described above.

Koenig uses utterance final creaky voice far more in the podcast than the interview. A chi-square test for independence shows that the relationship between the context and amount of utterance final creaky voice is significant, X^2 (1, N = 447) = 51.7, p < .00001, confirming my hypothesis that the amount of utterance final creaky voice is dependent on the context. Over the same length of speech, in the Fireside Interview, she exhibits 33% fewer instances of utterance final creaky voice than in *Serial*. While in the interview, the majority of her utterances end without creaky voice, in *Serial*, the majority end with it.

With Sarah Koenig's history as a journalist and reporter, and the nature of her work, she might be trying to index similar things to newscasters. As described above, the key characteristics of the journalistic persona that Koenig is accessing with her creaky voice are authority and engagement. Her use of more creaky voice in the podcast would suggest that she uses her journalistic persona when conveying facts and telling a news story. In the podcast, she is conveying facts, while in the interview she talks about her personal opinions and experiences. Her use of creaky voice lines up with our hypothesis that she would have larger excursions in lower boundary tones to convey authority, as described by Gasser et al. (2019).

4.1 Fireside Interview

A closer look at the instances of Koenig's creaky voice in the interview shows that she uses the feature to access an authoritative journalistic persona. In the quotes below, bold italicised words have creaky voice and slashes indicate utterance boundaries. The quotes are in response to the question: "As a journalist, maybe just talk a little bit about, what is your product?".

(1) No, I really, I have a very very old-fashioned / notion of what it is that I ***do***.
(2) And they're trusting me to like / then go put it out in the ***world***. That's kinda how I think ***about it***.
(3) That is a big, big ***responsibility*** / to have people trust you ***with that.***
(4) I wouldn't trust anybody with that / So, I take it really ***seriously.***
(5) It's my job to shine a light / on stuff that / is wrong, ***basically***.
(6) You cannot be an advocate for ***anything***, / you can't be ***an advocate***.
(7) That's not what I'm ***doing***, / I'm not an ***advocate***.
(8) If you have researched something thoroughly ***enough*** / and rigorously ***enough***, / I think it's ok to take a ***position***.
(9) So, it was in a way more ***reporterly***, / or more traditional ***reporterly***.

The above quotes demonstrate that Koenig uses creaky voice in her speech to index seriousness. The words that coincide with her creaky voice are those that are focused on being emphatic about the job of a journalist or semantically related to journalism and reporting. She takes a stance on the role of journalists and the power they have over people's stories, describes what it means to be "reporterly" or "traditional reporterly", and differentiates "advocating" from taking a position. It would seem that in the interview context, creaky voice is a feature of Koenig's speech that she accesses to index journalistic authority. These findings are in line with Podesva's (2007) findings about the use of prosodic features as a means of constructing linguistic identities. So, we might describe Koenig's use of creaky voice as constructing a journalistic identity. The three sentences below further showcase how Koenig shifts into this persona, as well as what is specifically indexed when this persona is accessed.

(10) So, you have to sort of, be aware of your biases, / try to set them aside and be like, / "Just teach me, / I'm here to learn, / just teach me, teach me".
(11) But then you get to the end of that road / and you're like, / "Nah, there's like, / there's an up or down here / and like, I think it's okay to say like, / nah or nah".
(12) You really have to do your homework ***to get there*** / and to be trusted by the ***audience***.

In (10), she is indicating that part of a reporter's job is to put biases aside. In (11), the trajectory of her thought shifts as she describes times where, as a reporter, she does have to make moral judgements. Finally, in (12), she states how much work has to be done before the point where a reporter can take a position. In these examples, she uses quotative speech to communicate her point. There is no creaky voice until she is making a decisive statement about how difficult it is to get to that stage in reporting a story. So, her creaky voice may also specifically index more authority and knowledgeability, which is consistent with Gasser et al. (2019).

There are instances where creaky voice appears, but the context is not especially related to journalism. This can be accounted for considering that Koenig has been invited to speak authoritatively on her work, so she would be trying to index authority throughout all her speech. An additional alternate account for Koenig's creaky voice is the interactional purpose of the feature. When looking at the sentences, creaky voice seems to only occur at the end of a completed thought. In sentences (10) – (12), she only uses creaky voice at the end of her anecdote. So, creaky voice may serve the purpose of letting her interlocuter (the interviewer) know she is done speaking. However, her increased use of creaky voice in the podcast context shows that it is not just an interactional feature.

4.2 *Serial* Podcast

Koenig's use of creaky voice in the interview context can be contrasted with the distribution of creaky voice in the *Serial* podcast. In this context, there is no interviewer and she is telling a prewritten story. The quotes below show examples of her creaky voice distribution in podcast utterances.

(13) About a year and a half ***ago***, / clips from this ***video*** / appeared on every major news ***broadcast***.
(14) It showed the rescue of a guy named Bowe ***Bergdahl,*** / he was the U.S. soldier who was captured by the Taliban / and held captive for just shy of ***five years.***
(15) In the backseat of the truck / a door is ***open***.
(16) Even ***now***, / other soldiers are so enraged by what Bergdahl did / that for his own protection / he's got to have a security detail ***with him*** / when he leaves his base in San Antonio.
(17) And through all ***of this***, / Bowe Bergdahl himself was like ***a ghost***, / a ***blank***.
(18) He wasn't talking on TV, / he wasn't quoted in the ***newspaper***, / so it seemed like that was ***that***.
(19) But last spring / I found out that Bowe Bergdahl / had been talking ***to someone*** / for almost a ***year***.
(20) He'd been talking to a film ***maker*** / named Mark ***Boal***.

In addition to being more frequent, her use of creaky voice seems to not be correlated with any particular topic. For example, in sentence (18) where we might expect her to use creaky voice consistently while listing, there is creaky voice on the word "newspaper" but not on the word "TV". This would suggest that Koenig's use of creaky voice in the contexts is different. She uses creaky voice in her interview to index authority and reinforce her journalistic identity, whereas in this context, it is a fully integrated part of her speech, because when she is reporting, she needs to index authority and trustworthiness in everything she says, regardless of content. In this context, the utterances without creaky voice were typically in the middle of a sentence or thought; however, there does not seem to be any particular pattern. Noncreaky voice utterances seem to be arbitrarily dispersed in a similar way to utterances with creaky voice.

If it is the case that Sarah Koenig's creaky voice is something that has increased over time as Liberman (2015) finds, it seems to be a prevalent feature in all contexts. Ultimately, when comparing the two contexts, we see that Koenig uses more creaky voice in the podcast context than the interview context, which suggests that she is style-shifting. Her increased use of utterance final creaky voice in *Serial* agrees with previous literature about newscaster prosody. This usage indicates that Koenig is meaningfully using creaky voice as a way to tell a story while indexing trustworthiness and engagement, as detailed in Gasser et al. (2019). Additionally, in the interview context, her use of creaky voice is often specific to utterances meant to index authority, which could be her tapping into that journalistic identity used in the podcast. This is evident in her use of utterance final creaky voice as a focusing technique when talking about the job of a reporter and using words semantically related to journalism in the interview. From these observations, we can say that Koenig is constructing an authoritative journalistic identity with her use of utterance final creaky voice.

5 Discussion

From this case study, we see that from a speaker design perspective, creaky voice can be an effective tool in constructing speaker identity. The previously discussed studies of creaky voice have focused primarily on listener perception. While more research is needed to make a claim about whether listeners would judge Koenig's creaky voice favourably, this study offers us insight into how creaky voice may be used by the speaker to create specific identities. With regard to previous analyses of Koenig's increase in creaky voice, this change over time might indicate an evolution in Koenig's confidence as a journalistic authority leading to increased comfort in using and accessing this journalistic identity.

This study is limited in its lack of control variables. There were no matching sentences between the two contexts, and the environment of the contexts differs greatly. Additionally, with the presence of an interlocutor, it is ambiguous whether Koenig's creaky voice in the interview context is for constructing identity or for interactive purposes. Looking at a context without a live audience would be one way to minimise the differences between contexts, however this does not eliminate the interlocutor issue.

When considered in the context of past studies on the social meaning of creaky voice, this study explores an example of a woman using creaky voice to index the authority, trustworthiness, and engagement needed by a reporter. Yuasa (2010) discusses how in the past, creaky voice in men has been a feature seen as a marker of confidence, authority, and masculinity, and suggests that metropolitan women may be accessing the feature to similar effect. Koenig might be an example of how the "authoritative" social meaning of creaky voice has become a feature for women as well. Yuasa (2010) also describes how American female speakers use creaky voice regularly regardless of conversational context; however, the findings in this paper may indicate that there are particular contexts where creaky voice is more common. The findings here might also corroborate Gasser et al. (2019), who connect lower L% tones with authority. This study also builds on the discussion of how semantic context can be a useful way to identify the social meaning of a feature, as in Podesva (2007).

For further study, looking at the differences between the first and second seasons of *Serial* might provide more insight into Koenig's various personae and how they might affect her voice. As Koenig noted that her personal connection to the story in the first season might have altered her voice, there may be measurable differences between her "girl sleuth" persona and her "journalistic" persona. In particular, the findings of this study might indicate that Koenig might exhibit a wide variety of linguistic features in many more identities than the ones I have discussed.

Beyond that, since the release and popularity of *Serial*, there have been many other podcasts in the same "true crime" genre. Studies into the perception of podcaster speech and differentiation between podcaster and newscaster prosody would also be interesting to explore. From further studies, we might be able to determine differences between what podcasters (in this particular podcast genre) and newscasters index with their linguistic features.

Looking at other similar podcasts and podcast hosts would allow us to make more generalisations about "podcaster prosody". Because *Serial* is the most downloaded podcast of all time, it would stand to reason that many other presenters looking to be successful in the medium might emulate what Koenig is doing. What we might find is deviations from the prosodic features of traditional newscaster speech to index a new set of features. As Eckert (2020) describes, young people seem to find more personal and informal presentation styles more engaging, so features that index youth and authority like creaky voice may become a wider trend for news presenters and podcasters. Based on Koenig's own descriptions of her *Serial* persona, the goals of a "podcaster" persona may differ from the "journalist" persona in several ways. What remains to be investigated is how the rules of journalism and reporting are changing because of podcasts, how that might be reflected in a change in what it means to be a journalist, and how that change might finally be reflected in linguistic features. If we identify Koenig's "girl sleuth" persona from the first season of *Serial* as more indicative of the wider "podcaster" trend, then we may find that podcasters (at least those in the same genre as *Serial*) are more comfortable with uncertainty and may aim to engage the listener on a more personal, intimate level. Mottram (2016) suggests that it may be futile to attempt to describe trends in how podcasters construct their identity. She writes that "in podcasting, achieving one's 'vocal authority' means finding one's individual authenticity" (Mottram 2016:53). This may mean that there could potentially be fewer commonalities between podcasters than between newscasters, as each podcaster tries to set themselves apart as an individual. In any case, the medium of podcasting is still relatively new, and the evolution of podcast hosts and their language goals is something to observe through future linguistic study.

Acknowledgements

I would like to thank Dr. Maya Abtahian at the University of Rochester for all her help on this project.

References

Anderson, Rindy C., Casey A. Klofstad, William J. Mayew, and Mohan Venkatachalam. 2014. Vocal fry may undermine the success of young women in the labor market. *PLoS ONE* 9(5):e97506. https://doi.org/10.1371/journal.pone.0097506

Barclay, Steven. 2020. Sarah Koenig [Online biography]. Accessed 23 August 2020, www.barclayagency.com/speakers/sarah-koenig/

Boersma, Paul. 2001. Praat 6.1.09 [Software]. www.praat.org/

Coupland, Nikolas. 2002. Language, situation, and the relational self: Theorizing dialect-style in sociolinguistics. In *Style and Sociolinguistic Variation*, ed. P. Eckert and J. Rickford, 185–210. Cambridge, UK: Cambridge University Press.

Eckert, Penelope. 2011. Language and power in the preadolescent heterosexual market. *American Speech* 86(1):85–97. https://doi.org/10.1215/00031283-1277528

Eckert, Penelope. 2020. Indexical obsolescence. In *The Routledge Companion to the Work of John R. Rickford*, ed. R. Blake and I. Buchstaller, 298–307. New York and London: Routledge.

Gasser, Emily, Byron Ahn, Donna Jo Napoli, and Z. L. Zhou. 2019. Production, perception, and communicative goals of American newscaster speech. *Language in Society* 48(2): 233–259. https://doi.org/10.1017/S0047404518001392

Higdon, Michael J. 2016. Oral advocacy and vocal fry: The unseemly, sexist side of nonverbal persuasion. *Legal Communication & Rhetoric: JALWD* 13:209–219. https://www.alwd.org/index.php?option=com_attachments&task=download&id=50

Koenig, Sarah. 2015. Serial: Dustwun [Podcast episode]. Accessed 23 August 2020, https://serialpodcast.org/season-two/1/dustwun

Liberman, Mark. 2015. Sarah Koenig [Blog post]. Accessed 23 August 2020, languagelog.ldc.upenn.edu/nll/?p=17549

Ligon, Claire, Carrie Rountrey, Noopur Vaidya Rank, Michael Hull, and Aliaa Khidr. 2018. Perceived desirability of vocal fry among female speech communication disorders graduate students. *Journal of Voice* 33(5): P805.e21–805.e35. https://doi.org/10.1016/j.jvoice.2018.03.010

Mottram, Christine. 2016. Finding a pitch that resonates: An examination of gender and vocal authority in podcasting. *Voice and Speech Review* 10(1):53–69. https://doi.org/10.1080/23268263.2017.1282683

Parker, Michelle A., and Stephanie A. Borrie. 2017. Judgments of intelligence and likability of young adult female speakers of American English: The influence of vocal fry and the surrounding acoustic-prosodic context. *Journal of Voice* 32(5):538–545. https://doi.org/10.1016/j.jvoice.2017.08.002

Podesva, Robert J. 2007. Phonation type as a stylistic variable: The use of falsetto in constructing a persona. *Journal of Sociolinguistics* 11(4):478–504. https://doi.org/10.1111/j.1467-9841.2007.00334.x

Point Park University. 2016. Sarah Koenig on her journalistic voice in the Serial Podcast [YouTube video]. Accessed 23 August 2020, www.youtube.com/watch?v=JlyQgwFVlNk

Quah, Nicholas. 2018. Serial Season 3 is the podcast's biggest ever [Online article]. www.vulture.com/2018/12/serial-season-3-50-milllion-downloads.html

Qualtrics XM. 2017, 28 Oct. Serial Fireside Chat - Sarah Koenig [YouTube video]. Accessed 23 August 2020, www.youtube.com/watch?v=xlmbtaEg1X8

Rodero, Emma. 2014. The principle of distinctive and contrastive coherence of prosody in radio news: An analysis of perception and recognition. *Journal of Nonverbal Behavior* 39(1):79–92. https://doi.org/10.1007/s10919-014-0201-5

Roose, Kevin. 2014. What's behind the great podcast renaissance? [Article]. https://nymag.com/intelligencer/2014/10/whats-behind-the-great-podcast-renaissance.html

Ward, Nigel. 2006. Non-lexical conversation sounds in American English. *Pragmatics and Cognition* 14(1): 120–182. http://doi.org/10.1075/PC.14.1.08WAR

Wolf, Naomi. 2015. Young women, give up the vocal fry and reclaim your strong female voice [Article in *The Guardian*]. https://www.theguardian.com/commentisfree/2015/jul/24/vocal-fry-strong-female-voice

Wolk, Lesley, Nassima B. Abdelli-Beruh, and Dianne Slavin. 2012. Habitual use of vocal fry in young adult female speakers. *Journal of Voice* 26(3):e111–e116. https://doi.org/10.1016/j.jvoice.2011.04.007

Yuasa, Ikuko Patricia. 2010. Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women? *American Speech* 85(3):315–337. https://doi.org/10.1215/00031283-2010-018

Zhang, Qing. 2005. A Chinese yuppie in Beijing: Phonological variation and the construction of a new professional identity. *Language in Society* 34(3):431–466. https://doi.org/10.1017/s0047404505050153

eyau3@u.rochester.edu