NORTHWESTERN UNIVERSITY

Social Expectations in Linguistic Memory

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Linguistics

By

Amelia Stecker

EVANSTON, ILLINOIS

October 2023

© Copyright by Amelia Stecker, 2023

All Rights Reserved

Abstract

Listeners engage social knowledge and attitudes about speakers in processes of speech perception (Niedzielski 1999; Drager 2011; Levon 2014), and correspondingly construct and modify cognitive representations of sociolinguistic information (Sumner et al. 2014). However, relatively little is known about the structure of sociolinguistic representations and their relation to cognitive processes. This dissertation explores the social factors that influence linguistic memory as a way to understand how sociolinguistic representations are formed in the mind.

While sociolinguistic memory remains largely understudied, some experimental work has illustrated that speech recognition is biased to support listeners' pre-existing social expectations (D'Onofrio 2021), illustrating the ways in which social ideologies and attitudes guide how listeners 1) remember information in speech and 2) recognize the use of a socially coded prosodic contour, i.e. uptalk. Specifically, my dissertation explores the influence of gender ideologies on sociolinguistic memory, as instances of metalinguistic commentary reveal biases against women's voices and the linguistic features associated with them (Gross 2015).

In three experiments, I analyzed listeners' recall of speech content and recognition of speakers' use of falling and rising utterances, the latter of which could be interpreted as tokens of uptalk. In Study 1, I assessed whether participants' memory for a speaker's use of rising and falling utterances were conditioned by the speaker's perceived gender. Study 2 tested the extent to which listeners falsely recognized rising utterances when this contour was in fact never produced by their speaker, investigating rates at which listeners falsely recognized rising tokens when listening to speakers perceived categorically as men versus women. In Study 3, I then analyzed whether the patterns exhibited in Study 1 persisted when listeners were primed with brief metalinguistic commentary on uptalk. Exploratory analyses following the main experiments

illustrated that listeners' attitudes regarding standard language ideology and uptalk modulated the rate at which they recognize falling and rising utterances. Listeners' performance across these tasks revealed the ways in which ideological factors, including listeners' attitudes and the perceived gender of the speaker, influence what linguistic information is stored in memory and thus cognitively represented.

Acknowledgements

It's a researcher's dream to answer the questions that stuck out to you as a kid. I'm incredibly grateful for support from the Linguistics Department, friends, and family I've received in my pursuit of this dissertation, which stems from my curiosity of how our perceptions of language reflect gendered biases – something I've felt from the first instances of being told "not to sound like *that* kind of girl."

I'm first and foremost thankful to my primary advisor, Annette D'Onofrio. It's been a blessing to be one of your first students, to be welcomed and appreciated as my authentic self in a scientific world that, if not pushed back, can feel dehumanizing to the researcher and their work. While training me as a sociolinguist, you taught me how to articulate exactly what I want people to hear; how to imbue research with the humanity that calls for these questions to be answered in the first place. Most of all, thank you for your model of prioritizing health over school – your compassion for mine and others' wellbeing has been unwavering, especially in the times of a global pandemic.

Thank you to my committee members Jennifer Cole and Gregory Ward for your insightful feedback on this project and for inspirational discussions inside and out of the classroom. I'm grateful for your expertise in your respective fields as your guidance so very much enriched this project and its theoretical implications in the worlds of prosody and gender.

I'm very grateful for other systems of support at Northwestern University: This dissertation was financially supported by Graduate Research Grants from The Graduate School and the Department of Linguistics. I'm also grateful for members of SocioGroup for their consistently insightful feedback upon hearing the multiple renditions of this project. Thank you to Chun Chan for your support in all things technical, including everything in between from

virtually launching my experiments to turning on my mic for my public defense. I feel quite fortunate as a graduate student for your steadfast attention to troubleshooting any logistical issue and for your patience in communicating the solution in numerous back-and-forth email chains at any time of day.

Thank you to my lovely cohort-mates Jaime Benheim, Maria Gavino, and Wes Orth for feeling like family whenever we studied for too long in the cubicles, grabbed drinks at Bangers and Lace, or Zoomed in from our parents' homes during the pandemic. I probably couldn't have come out the other side of the learning curve that is statistics and learning to code without you all, and I'm so much more grateful to have been surrounded by peers who work with such inspirational curiosity. Additionally, thank you to Tommy Denby, Jordan Schulte, Jennifer Dibbern, Shawn Foster, Lisa Cox, Chantal De Leon, Anna Robinson, Kate Sandberg, Thomas Sostarics, Cassie Davenport, and many others for making these Chicago winters feel warm with your camaraderie.

Thank you to my Prolific participants who, according to some responses, may not have enjoyed the daunting task of having to remember what was said and how the speaker said it. I'm grateful to have had sampling access through this platform to all kinds of people, all of with whom I would have never interacted if it wasn't for their interest in taking my study from wherever they are in the world.

To my unlucky friends and family whom I recruited to record stimuli, thank you for allowing me to stuff you in a closet and make you repeat the same story over and over. These sessions gave me two lessons: first, uptalk is hard to imitate! And unless you can emulate (or you are) a hyper twelve-year-old kid excitedly telling a story about their day, it feels quite unnatural to use in one sentence after the other. And second, my friends are a patient bunch who wanted

me to answer exactly the questions about the world that I couldn't stop asking. I can't wait to repay you for your services in baked goods and hugs.

As a Cancer Sun, I think I could write a separate dissertation on the love and support that I've been so fortunate to have as long as I've been both a person and a student. To my mom Betsy, I've loved language since you actually taught me as your student with root word and spelling tests. To my dad Tom, I potentially feel most proud of myself when I think about you living as a grad student on Orrington – I love being the daughter of an architect, a thinker. David, you've always been proud of me as your big sister even if I never really ever told you what I was doing at grad school. Thank you all for making me feel loved every day regardless of my academic accomplishments, and for framing "getting a linguistics doctorate" as just one of the many things I'm capable of and one of many meaningful ways I can contribute to the world.

Finally, to my husband Noah, it's a gift to do anything with you – even if it's writing a dissertation while you're sleeping next to me as I give in to a second wind. There's something about a latte made by you that makes being industrious gratifying again. I love all your answers to when I ask you, "What's that word for when…," hoping that you just might know exactly what I'm trying to say. I promise your suggestions have been helpful more than a couple times. When I started this program, I had thought that it would push back a wedding; our excitement for starting life together made that just not so. And as a monolingual linguist, it's a dream to love someone who loves learning languages almost as much as they love you. Thank you for everything.

Table of Contents

List of tables

List of figures

1. Introduction and literature review

1.1. Introduction

As native speakers of a language, individuals store sociolinguistic elements of incoming speech in memory (Sumner, Kim, King, & McGowan 2014; Pierrehumbert 2016), and this encoding aids in further recall and recognition in future processes of perceiving and interpreting new information from speech (Babel & Russell 2015; McGowan 2015; Squires 2016). In particular, listeners' performance in speech perception tasks such as phoneme categorization and word identification can depend on their intake of social information, such as gender or age, visually or acoustically cued by the speaker (Mann & Repp 1981; Koops, Gentry, & Pantos 2008). Listeners' mental mappings of their world therefore call upon and uphold cognitive representations of socially meaningful language and social categories of speakers who produce it (Drager & Kirtley 2016). Listeners' prior experiences of perceiving and interpreting variation in linguistic stimuli, and subsequent learning from these experiences, drive how they take in new linguistic information from new or previously known speakers (Bradlow, Nygaard, & Pisoni 1999; Church & Schacter 1994; Goldinger 1996).

Social psychological work has found that existing structures in individuals' memories influence how they perceive incoming information. In examples of *confirmation bias,* listeners are more likely to seek evidence for their existing beliefs by retaining information that aligns with their beliefs, and ignoring information that opposes these existing frames (Wason, 1960; Taylor & Brown 1988; Nickerson, 1998; Jonas, Shulz-Hardt, Frey, & Thelen 2001; Oswald & Grosjean 2004). Listeners' responses to new incoming information therefore have been shown to exhibit *congeniality effects*: via motivated reasoning (Kunda 1990), individuals generate and evaluate causal theories – what was said, by whom, and why – in a self-serving manner (Kunda

1987). The question arises: Do listeners attend to socially meaningful linguistic productions in the same way?

Existing sociolinguistic empirical and theoretical work has posited that listeners' existing social knowledge of their world co-operates with their mental encoding of linguistic input (Foulkes & Docherty 2006; 2010; Sumner et al. 2014; D'Onofrio 2021). Speech perception processes such as phoneme classification and listeners' evaluations of their speaker are guided by social associations cued in the signal (Johnson, Strand, & D'Imperio 1999; Campbell-Kibler 2007; Levon 2007; D'Onofrio 2018) and its interaction with top-down overt priming of social information (Strand 1999; Niedzielski 1999; Hay, Nolan & Drager 2006; Squires 2013; D'Onofrio 2018).

To theorize this within exemplar models of linguistic cognition (Johnson 1997), listeners encode linguistic instances from external speech as exemplars, mental episodes encompassing specific information of the acoustic signal and context, and store them according to dimensions of perceptual similarity (Pierrehumbert 2001). Listeners have been found to retain speaker-specific information that aids in future speech processing (Goldinger 1996; Staum Casasanto 2009). These results provide evidence that listeners' mental conceptions of socially structured linguistic variation play a role in guiding ideologically biased sociolinguistic perception. Listeners theoretically alter their attentional weight and cognitive space for specific linguistic forms that are associated with existing social categories (Sumner 2015). However, less attention has been assigned to exploring the exact nature of these social expectations as factors in sociolinguistic cognition: where they come from, and the degree to which listeners exhibit the aforementioned cognitive biases, such as confirmation bias and congeniality effect phenomena, in their linguistic memory.

At varying levels of awareness (D'Onofrio 2018; Labov, Ash, Ravindranath. Weldon, Baranowski, & Nagy 2011; Campbell-Kibler 2016; Carmichael 2016; Sharma 2018), listeners and speakers interactively hone their knowledge of social semiotic structure in language – the social indexicality of linguistic features (Eckert 2008, 2012). Individuals' linguistic memory processes underlie language perception in practice (Drager 2010). For example, Hay, Nolan, & Drager (2006) studied perceptions of a New Zealand and Australian English phonetic distinction. When participants were primed with social information that categorized the speaker as Australian, listeners were more likely to hear a higher and fronter /ɪ/ vowel than when the speaker was categorized as a New Zealander. In this way, these participants provided evidence of the activation of particular linguistic exemplars that are linked to social categories, like speakers of a particular dialect. While this entails that listeners' ideological understandings of different social groups or types of speakers contribute to the storage and recollection of linguistic information, less is known about what exactly comprises these social expectations and activated links between social categories and linguistic episodes that are at work in memory. More empirical attention is required to understand the ways in which listeners' cognitive representations are sensitive to the social and stereotypical meanings couched in and reflected by linguistic variation.

In a series of memory tasks, this dissertation empirically tests listeners' recognition of utterances with falling and rising final contours, the latter of which can be perceived as tokens of the psycho-social linguistic phenomenon *uptalk.* Speakers use uptalk within larger linguistic styles (Podesva & Kajino 2014) for floor-holding, listener confirmation, and a variety of other discourse functions (Innes 2007; Guy & Vonwiller 1984; Allen 1984; Levon 2016), but popular commentary and early academic attention to this rising tone stereotypically associate its use with

expressions of femininity (Green 2014; Lakoff 1975). While work in linguistic production has suggested that women are more likely to incorporate uptalk into their linguistic styles (Linneman 2013), speakers of all genders use uptalk in different contexts and for discursive listener-sensitive purposes (Levon 2016; Warren 2016). By no coincidence, uptalk's indexical links to femininity coincide with negatively evaluated attributes such as sounding "ditzy," "vapid," or "insecure" (McConnell-Ginet 1983). In instantiations of metalinguistic commentary, interpretations and consequences of using uptalk have been described as "examples of how fraught it can be to speak while female" (Romm 2017). How might gendered biases against women's voices and social meanings of uptalk as a gender-linked linguistic feature manifest in explicit recognition memory and underlying cognitive representations of speech?

In three experiments, my dissertation engages an existing framework that centers listeners' memory as a cognitive site for the intake and representation of linguistic knowledge as well as the maintenance of mappings of language to the social world (Hay & Drager 2010; Tripp & Munson 2021). I delve into gender as an ideological construct that productively regulates attitudes and expectations about speakers' voices to investigate the following research question: Do listeners' expectations and existing stereotypes about gender modulate their memory of speakers' speech content and use of sociolinguistic features?

In Study 1 (Chapter 2), I test listeners' accuracy of their speech content recall and their old-new recognition of rising utterances previously categorized as uptalk (e.g. Warren 2016; Tyler 2015). In Study 2 (Chapter 3), I examine listeners' rates of falsely remembering instances of rises, in an experimental context in which these utterances were never actually produced by their speaker. Study 3 (Chapter 4) examines a stereotypical prime's effect on participants' memory by presenting participants with metalinguistic commentary regarding social

expectations about which gendered types of speakers use uptalk most before asking participants to perform the same task as in Studies 1 and 2 (Chapters 2 and 3). Finally, I explore specific social and demographic factors from listeners' backgrounds and examine their effects on old-new recognition of rising and falling utterances.

Results suggest that listeners' (false) recognition of rising utterances involves salient links between uptalk as a prosodic phenomenon and female speakers as a gendered category, yielding greater rates of recognition for women's rises in contrast to men's rises. In this way, listeners' responses illustrate that what is encoded into and recalled from their working linguistic memory is complicated by existing stereotypical expectations of uptalk and gendered voices. Building on emerging sociolinguistic work that centers social meaning in linguistic practice (Eckert 2012; Eckert & McConnell-Ginet 2003; Zhang 2021) and more specifically sociolinguistic perception (D'Onofrio 2021; Drager, Hardeman-Guthrie, Schutz, & Chik 2021), this dissertation posits that listeners' existing expectations about their speaker and gendered stereotypes mediate their cognitive storage of speech.

1.2. Sociolinguistic perception and memory

Listeners' social expectations about a speaker inform processes in their speech perception (Strand 1999; Niedzielski 1999; Koops et al. 2008; Drager 2011; Campbell-Kibler 2006, 2007; D'Onofrio 2018). Previous work has found that listeners are sensitive to acoustic cues that index (Agha 1993; Silverstein 2003; Eckert 2008, 2012) different social types of speakers (Podesva 2011). For example, Strand (1999) explored the ways in which gendered expectations of voices constrain low-level linguistic perception. In a voice-only experiment, listeners' categorization of sibilant tokens along an /s/–/ʃ/ continuum was guided by the gendered information they retrieved

from the rest of the voice – other socially meaningful cues, such as fundamental frequency and pitch information, that prompted the listener to assume whether they were listening to a male or female voice. Categorization of sibilant tokens followed gendered lines, even though differences in sibilant frequency is not derived from any physiological differences between these genders (Fuchs & Toda 2010). When listeners heard a voice they perceived as male, their categorization of /s/ tokens comprised a larger range of acoustic frequencies in contrast to a more confined range of higher frequencies for /s/ tokens from a speaker perceived as female. Work like Strand's (1999) study provides a foundation for exploring the connection between social expectations of voices, particularly about gendered speakers and speech perception.

Other work has expanded on the impact of different types of social information on a variety of processes in speech perception. For example, social information prompting a speaker's gender identity (Johnson, Strand, & D'Imperio 1999; Strand 1999; Winn, Tripp, & Munson 2013; Munson, Ryherd, & Kemper 2017; Alderton 2020), age (Harrington, Kleber & Reubold 2007, 2008; Drager 2011), geographic region (Niedzielski 1999; Hay, Nolan & Drager 2006), socioeconomic background (Hay, Warren & Drager 2006; Squires 2014; Lawrence 2017), racialized identity and perceived ethnicity (Rubin 1992; Yi, Phelps, Smiljanic, & Chandrasekaran 2013; McGowan 2015; D'Onofrio 2019), and social-type persona configuration (D'Onofrio 2018; Villareal 2018) have been found to condition listeners' perception at different linguistic levels (from phoneme categorization conducted by Strand [1999] to grammaticality judgments found by Squires [2013]) and at a different degrees of awareness (Labov et al. 2011; Levon & Fox 2014). In this dissertation, I build on these established findings that suggest an influence of social information on processes of linguistic cognition, such as speech perception and recognition (Thomas 2002; Sumner et al. 2014; Drager 2015; Drager & Kirtley 2016).

In an exploration of sociolinguistic perception and cognition, this dissertation follows an emerging theme in sociolinguistic research that foregrounds the role of the listener in processes of social meaning-making through language (Levon 2014, Campbell-Kibler 2008). Listeners, constrained by their own sociolinguistic experiences and knowledge, often interpret a speaker's behavior in ways that the speaker may not intend. Levon's (2014) experimental findings illustrate that listeners' endorsements of sexuality and gender stereotypes modulated how listeners evaluated speakers' use of linguistic variables linked with gender identity and sexual orientation. In this way, as listeners perceive an incoming speech signal, the ways in which they attend to and construct social meaning embedded in speech is guided by their own ideologies and expectations of others (Levon 2014). Here, I define *ideology* as a meaning-making framework that guides individuals' interpretations of and movements within their social world or public sphere (Ochs 1992, Irvine & Gal 2000). Ideologies direct the construction and maintenance of beliefs and attitudes (Milroy 2004), which individuals may apply to external stimuli, such as other individuals or incoming information (Huskinson & Haddock 2006; Eagly & Chaiken 1993; Zanna & Rempel 1988). Sociolinguistic research has predominantly engaged with individuals' ideological standpoints by exploring the ways in which speakers' dynamic social orientations to their community underlie sociophonetic change and variation (ex. variation in speakers' acoustic space of diphthongs /ai/ and /au/ in Martha's Vineyard [Labov 1963]; Chicagoans' reversal of the Northern Cities Vowel Shift [D'Onofrio & Benheim 2019]; lowering of the TRAP vowel, often referred to as 'short-a', in acoustic space by different racialized and persona-typed speakers in Rochester, NY [King 2021]). However, relatively less scholastic attention has been allotted to understanding how individuals' ideological bearings, such as their attitudes and expectations of others, guide the inner workings of their own linguistic memory and cognitive faculties.

An understanding of memory and the ways in which listeners attend to, encode, and recall particular pieces of information can illustrate how listeners build cognitive representations to facilitate learning and other tasks (Squire 2004), such as sociolinguistic perception. Research has shown that when processing speech streams, listeners attend to and store tokens of linguistic utterances in fine contextual and phonological detail (Church & Schacter 1994; Nygaard, Sommers & Pisoni 1994). Investigations of linguistic cognition illustrate that listeners' attention to and memory of utterances are mutually dependent: Individuals' attention allotments are biased to uphold existing memory structures (Chun & Turk-Browne 2007) and memory of particular objects (such as linguistic forms or speakers) is conditioned by the attention that an individual provides to that form (Bunting, Cowan, & Colflesh 2008). Understanding the nuanced role of attention in sociolinguistic perception has proven to be a difficult task. Sociolinguistic research has predominantly focused on attention as a factor in variation in linguistic production. For example, in the attention-paid-to-speech model (Labov 1972), listeners are theorized to use stigmatized linguistic features less frequently when they pay more attention to their own speech.

As the encoding and retention of information relies on the individual's maintenance of attention to incoming stimuli (Baddeley 2003), some experimental work has operationalized attention in listeners' sociolinguistic perception. For example, Austen & Campbell-Kibler (2022) tested listeners' real-time evaluations of speakers producing stimuli containing socially meaningful linguistic features. Listeners' social evaluations were predicted to be modulated by the online noticing and attention to the speaker's production of sociolinguistic features, such as (ING) and discourse marker *like*, as well as gestalt style-shifting between sounding "disfluent", "normal", "enthusiastic", or "bored" (Austen & Campbell 2022, p. e111). Listeners' responses indicated that their social ratings of their speaker exhibited poor time granularity and high

variability. These results suggest that individuals' attention to the production of specific linguistic features appears to interact with incoming stimuli in a more complex way than immediate reactions to isolated instances of socially salient productions. In this way, attention to and memory of speech are intertwined (Oakley 2009; Robinson 2003; Norman 1968): Attention guides what is stored in working memory, while existing knowledge structures in long-term memory simultaneously drive listeners' attention to new information. This dissertation explores the ways in which listeners' recognition of intonational tunes in utterances, as artifacts of their linguistic memory, were impacted by other factors I operationalized in each of the main studies. While the memory tasks in the main studies measured whether listeners remembered intonational productions, their performance in recognition was dependent on the information that they attended to during exposure in training phases.

Regarding memory for prosodic information, experimental work has shown that individuals retain information from a speaker's mean fundamental frequency to aid in future word recognition and recall information (Church & Schacter 1994; Goldinger 1998), and they also display subjective preferences for particular intonational structures (Waterworth 2003). More recent work on prosodic memory illustrates that prosodic characteristics are important components of information that are encoded in memory and guide future perception (Roettger & Rimland 2020; D'Imperio 2000; Severijnen, Rutger Bosker, Piai, & McQueen 2021; Baddeley 1992). Listeners have been found to attend to details of pitch accent, contour shape, and duration qualities in speakers' productions of lexical items and sentences (Cutler, Dahan, & Van Donselaar 1997). For example, listeners' capacity to distinguish new items from old in Kimball & Cole's (2016) study was found to be modulated by the contour shape in the signal. In a Hungarian perceptual study of linguistic focus, Kaldy & Barbarczy (2021) found that units

produced with greater focus (produced with 'eradicating stress,' the most prominent stress placement in the sentence [Kornai, & Kálmán 1988]) garnered more attention, and that this effect led to greater memory for these sentences.

Previous experiments in listeners' memory of prosodic information have often used digitally manipulated utterances that are maximally controlled for specific pitch characteristics in order to investigate the degree to which listeners can recognize (Kimball, Cole, Dell, & Shattuck-Huffnagel 2015; Kimball & Cole 2016) or imitate (Cole & Shattuck-Hufnagel 2011) differences in intonational productions. This body of empirical work has allotted more attention to investigating listeners' encoding of fine-grained prosodic details, rather than their recognition of utterances with categorically contrastive intonational tunes. More research is thus required to uncover whether social stereotypes can mediate memories of specific intonational productions.

1.2.1. Exemplar models of linguistic memory

This dissertation builds on existing research that examines individuals' memory to investigate the nature of underlying cognitive structures that store and represent linguistic information (Church & Schacter 1994; Green et al. 1997; Nygaard & Pisoni, 1998; Bradlow et al., 1999). While sociolinguistic memory is largely understudied, some theoretical and experimental work illustrates the relationship between social factors and processes in which speech is stored and represented in the mind. Foundational literature (e.g. Pierrehumbert 2001) theorizes that listeners store episodes of events as exemplars, which are cognitive representations of the information embedded in the event. These exemplars are stored in clusters, or exemplar clouds, organized by perceptual similarity along both social and linguistic dimensions. Following exemplar-based models, listeners record all acoustic information that they perceive from the signal in memory (Kimball, et al. 2015). Words or other linguistic forms (sounds, voices) that are produced more

frequently will constitute more robust cognitive representations in contrast to those that are less frequently heard, as more exemplars are represented in the mind for the more frequent form (Pierrehumbert 2001; Ellis 2002; Abramowicz 2007).

Listeners' episodic memory for spoken words entails highly specific linguistic representations that derive from existing detailed representations of previously spoken words (Craik & Kirsner 1974; Palmeri et al.1993; Goldinger 1996). For example, Goldinger (1996) investigated factors that mediated listeners' recognition memory of different speakers' linguistic utterances. He found that when listeners heard a word that was produced by the same speaker who had produced it in a previous training session, listeners remembered this word more accurately than when the word was produced by an unknown speaker. This suggests that listeners store phonetic detail embedded in incoming speech rather than discarding such information in processing.

Additionally, Goldinger (1996) found that information within individual speakers' voices facilitated listeners' word recognition, in that listeners were more accurate in remembering words that were produced by speakers of the same sex. These findings illustrate that listeners engage with social information about speakers' voices in the ways that linguistic utterances are encoded and recalled from explicit memory. Listeners' recognition of speech, as a process of both implicit and explicit memory (McBride-Chang 1996; Lachs, McMichael & Pisoni 2000), engages social information about linguistic features and the speakers who produce them (Sumner et al. 2014), and is biased to support their pre-existing social expectations (D'Onofrio 2021). Freeman & Ambady (2011) propose a dynamic interactive model through which listeners' construal of their speaker is temporally flexible and influenced both by top-down information regarding "high-level cognitive states and stereotype activations on categorization" as well as

"bottom-up category interactions due to shared perceptual features" (Freeman & Ambady 2011:1). In this way, listeners' interpretation of a speaker malleably evolves through their processing of information from both top-down and bottom-up routes.

1.2.2. Social weighting of linguistic episodes in cognition

As described above, listeners exploit overt social information and their perceptions of social characteristics of a speaker to guide the mapping of an acoustic signal to distinct lexical items or phones (Niedzielski 1999; Strand 1999; Hay et al., 2006; Babel 2009; Staum Casasanto 2009; Hay & Drager 2010; Munson 2011). When the acoustic input from a speaker opposes listeners' expectations and interpretations of what this speaker "should sound like," rates of processing have been shown to be slower or significantly hindered (Koops et al. 2008; Rubin 1992). Conversely, when the cued social characteristics and subsequent expectations about these voices are aligned with the speech signal, listeners' mapping of the acoustic signal to lexical representations can be enhanced (McGowan 2011; Szakay et al. 2012).

Further, listeners can falsely remember those linguistic instances that match their pre-existing ideological expectations about that speaker (D'Onofrio 2021). Some empirical work has shown that listeners can construct sociolinguistic expectations about a speaker without any top-down assignment of social identity categories. As an example of perceptual activation from bottom-up information, Sumner & Kataoka (2013) conducted a semantic priming task and a false memory task to test the semantic activation of target words and the interaction with listeners' exposure to (non-)rhotic tokens of specific English dialects: General American, New York City, and Southern Standard British English. Listeners who were primed with New York City non-rhotic tokens were more likely to falsely recall non-rhotic tokens they had never heard before in contrast to British English tokens, suggesting that listeners' recognition of these forms were

potentially derived from their own linguistic expectations what speakers of these different dialects "should sound like."

As another example, in a study of implicit and explicit memory effects for different variants of word-final /t/, Sumner & Samuel (2005) found that words that were produced with a word-final released /t/, the less-frequent and more "idealized" and "canonical" variant, were attended to and remembered better than words with more frequently produced but non-idealized variants, such as unreleased or glottal variants of word-final /t/. Since listeners are sensitive to social and contextual information in perception (Koops et al. 2008; Drager 2011; D'Onofrio 2018), Sumner et al. (2014) theorized that existing social knowledge may strengthen the encoding of exemplars linked with particular social concepts in cognitive representations (such as standardness, in the case of released /t/), leading to a stronger activation of these exemplars in speech perception in spite of lower frequency of occurrence.

Sumner et al.'s (2014) theoretical discussion provides a sturdy foundation that introduces social factors into a cognitive model of linguistic representation. However, their analysis does not delve into what exactly these social factors are and where they come from – why does word-final released /t/ have "idealized" or "canonical" qualities, which then mediate the ways this linguistic feature is cognitively stored? In other words, while Sumner et al. (2014) suggest that social information stemming from linguistic representations guides how incoming tokens are encoded, less is known about what these social components are and how they influence listeners' memory and cognitive representations of linguistic forms. This dissertation examines how ideological influences, such as social attitudes and expectations, mediate speakers' and listeners' sociolinguistic representations that are at work in linguistic production and perception.

D'Onofrio's (2021) study provides further groundwork by exploring the ways in which top-down social expectations may modulate listeners' memory of speech. In a word recognition task, participants were presented with or without information that a speaker was described as a "Business Professional," which is a persona embodying socioeconomic wealth, social mobility, and supra-local orientation, and is linked to the backed-TRAP vocalic feature, i.e. a backing in the vowel space and a decrease in the frequency of the second formant in tokens of the TRAP vowel (D'Onofrio 2015, 2018). Not only were participants better at remembering backed-TRAP tokens when their expectations about a speaker were socially congruent with the use of backed-TRAP, but this congruence of social and linguistic information also led participants to falsely classify *novel* backed-TRAP tokens as "old". In this way, listeners' interpretations and memory of sociolinguistic features may stem from ideological beliefs about the links between particular speakers and linguistic features. Listeners' linguistic memory, therefore, not only retains information from those real-world events and episodes, but it is also swayed by existing expectation-based mappings of what listeners think was said. These findings illustrate that recognition, as a process of explicit memory, may be biased to support listeners' pre-existing sociolinguistic expectations. Furthermore, these findings suggest that the cognitive representations of linguistic features, such as backed TRAP, contain social expectations about who uses these features, thus shedding light on the structure and influence of social factors in cognitive linguistic representations, as proposed by Sumner et al. (2014).

These theoretical conclusions from D'Onofrio (2021) allude to the possibility that a listener's sociolinguistic expectations may guide how they attend to and remember linguistic information – "not all experiences have equal social significance, and ideology can govern which forms or styles are worth attending to or remembering, or, in other words, are most strongly

connected to their social meanings" (D'Onofrio 2016: 17). As described above, forms imbued with greater social importance for an individual may glean greater allotments of attention, which lead to more robust mental conceptualizations. However, the extent to which a listener attends to specific components of a linguistic signal is difficult to measure explicitly in an experimental paradigm (c.f. Heuer & Hallowell 2015; Berends, Brouwer, & Sprenger 2016; Káldi & Babarczy 2021). Memory for linguistic information, however, can be more readily tested via listeners' performance in recognition and recall tasks, in which individuals are asked to classify whether a stimulus is novel or old from previous exposure and to provide information they previously heard. My dissertation examines the relationship between expectations about speaker gender and memory of sociolinguistic features by testing correlations between the perceived gender of a speaker and listeners' memory for content in speech and production of a salient gender-stereotyped linguistic feature.

1.3. Social expectations in linguistic perception

Social expectations and attitudes are constantly active in interaction (Simmons & Prentice 2006) and guide the ways in which people engage with others and interpret information. Attitudes, as products of ideologies, can guide what is salient or dominant – what information or experiences are prioritized over others – and what ought to be attended to in cognitive processing (Oosterhoff et al. 2018). Moreover, as explored by Eagly et al. (1999) in a meta-analysis, people are shown to more accurately retain information that supports their existing attitudes of social issues, in contrast to information that challenges those evaluations, a pattern known as the *congeniality effect*. This cognitive effect is corroborated by social psychological investigations of

confirmation bias, in which speakers tend to seek and interpret information partial to and consistent with existing beliefs (Nickerson 1998; Darley & Gross 1993).

1.3.1. Influences of listener attitudes on linguistic perception

Listeners' individual attitudes and endorsements of social stereotypes not only influence how they generally remember and make sense of information (Nickerson 1993; Kunda 1990) but also how they perceive speech (Campbell-Kibler 2007; Levon 2014; Babel & Russell 2015) and linguistically accommodate (including patterns of both convergence and divergence) to interlocutors (Babel 2010; Bourhis and Giles 1977; Giles 1973; Giles et al. 1991). For example, listeners in Levon's (2014) study who were more likely to endorse normative stereotypes about masculinity and male gender roles exhibited greater sensitivity to pitch and sibilance as phonetic cues to the social notions of "homosexuality" and "gayness" in social evaluations of speakers. Participants who exhibited less normative views on gender and sexuality displayed a different relationship between their perceptions of these features and evaluations of the speaker. Drager & Kirtley (2016) and Drager et al. (2010) posit that the prevalence and directions of these attitudes bias the activation of particular exemplars that align with the framing and positioning of these existing beliefs and attitudinal states. In this way, listeners' mental constructions of what was said and the embedded social meanings within the signal uphold their expectations of what they think they would hear from that speaker.

Individuals' memory may also be impacted by an incongruent pairing between a speaker's social identity and a speech pattern that violates expectations of that speaker. This element of surprisal, when a listener hears a speaker produce a linguistic feature that they would not expect from them, may also play a role in how productions of a particular linguistic feature are stored, and consequentially, how that feature is mentally represented (van Berkum, Hagoort,

& Brown 1999; Fine, Jaeger, Farmer, & Qian 2013; Squires 2014). Listeners may exhibit congeniality or surprisal effects at varying degrees, directions, and levels of awareness. While less frequently applied to sociolinguistic investigations, listeners' varying degrees of 'surprisal', as a proxy of the relative salience or noticeability of new information, has been shown to induce costs in language processing when a listener's predictions are not confirmed by stimuli (Hale 2001; Ranganath & Rainer 2003; Zarcone et al. 2016).

1.3.2. Metalinguistic expectations

Metalinguistic commentary – or 'talk about talk' (Bogetić 2016) – about social groups of speakers facilitates the construction of sociolinguistic expectations and complicates performance and accuracy of sociolinguistic perception (McGowan & Babel 2019). Previous work has explored listeners' varying degrees of stereotypical associations and enregisterment of particular linguistic features with social categories (Johnstone et al. 2006). Linguistic features classified as *stereotypes* garner greatest metalinguistic space and awareness over those features classified as *markers* or *indicators*, which point to more implicit or less-conscious links to social types or macro-social categories of people (Labov 1972).

Listeners' metalinguistic awareness of the social mappings to certain voices and linguistic styles are evident in the way that they discuss this practice, which, in turn, complicates what listeners believe they perceived from a speaker (Niedzielski 1999; Campbell-Kibler 2009; Carmichael 2016; McGowan & Babel 2019). Experimental sociolinguistic work has shown a variety of degrees to which listeners are sensitive to primed social and linguistic information in perceptual processes. For example, Hay et al. (2006) illustrated that New Zealand listeners presented with an "Australian" primed label for their speaker perceived /ɪ/ tokens as acoustically higher than when they were presented with a "New Zealander" label, though they were presented

with the same acoustic stimuli. In this way, existing sociolinguistic work provides evidence that listeners' speech perception processes are sensitive to linguistic features reaching metalinguistic salience. However, current theories do not yet provide thorough explanation for the relationship between listener's metalinguistic awareness of linguistic stereotypes and how it confines or challenges speech recognition and recall.

Processes in linguistic innovation, change, and variation happen through the vehicles of listeners' and speakers' own cognitive concepts of language (Dodsworth 2008; Drager 2010; Chevrot, Drager, & Foulkes 2018). These mental representations in language, ever active in linguistic praxis, are fundamentally at work in listeners' explicit and implicit memory. A theoretical understanding of sociolinguistic memory is crucial for recognizing the ways in which speakers and listeners imbue language with social meaning, enriching theories of the social structures that underlie linguistic variation and change. As listeners play a central role in the co-construction of social meanings conveyed by speakers in linguistic productions (Campbell-Kibler 2007; Levon 2008), their existing ideological framing of expectations of speakers has been shown to alter how they perceive incoming information (Strand 1999). How exactly do social expectations that a listener has about their speaker and their use of a socially meaningful linguistic feature bias what they take away from linguistic interactions? Previous work, while providing a foundation for understanding social expectations in sociolinguistic perception (Campbell-Kibler 2016), has yet to more holistically address how ideological underpinnings of what listeners think specific categories of speakers "should sound like" are constructed and maintained in listeners' memories and conceptualizations of linguistic practice. The studies in this dissertation address some aspects of these questions by calling on listeners' memory

processes (recall and recognition) as a fundamental site for the influences of social knowledge and ideological expectations on their linguistic expectations and representations.

1.4. Gender ideology and linguistic practice

This dissertation explores the ways in which listeners' memory is influenced by their existing social expectations about a particular social group of speakers. I focus on ideologies that productively regulate norms, attitudes, and stances that speakers use to position themselves in relation to others in their social world: ideologies surrounding gender.

1.4.1. Language and gender as a dynamic social project

Ideologies are very much at work in enacting "appropriate" culturally specific ways to perform or index membership of a gendered group, and they are pervasive across different levels of sociolinguistic practice (McElhinny 2003; Eckert 1989a, 1989b). Emerging expectations about gender continue to shift and operate in linguistic production (Eckert 1989, 1989; Calder 2019; Linneman 2013; McLemore 1991; Mendoza-Denton 2011), perception (Strand & Johnson 1996; Strand 1999; Stecker & D'Onofrio 2022), and interpretation and evaluations (Slobe 2016, 2018; Hachimi 2016; Gross 2015; Morgan 2015) of speakers of different genders.

Gender scholars have argued that categorization of gender is a ubiquitous trait of socialization as human individuals (Angouri 2021). Gendered categories are thought to facilitate a 'short-hand' for individuals as they perceive and interpret others: these groupings simplify otherwise complex pictures of individuals' gender identity and are imbued with meaning circulating in sociopolitical contexts and popular media (Angouri 2021). In this way, gender categories reflect idealized prototypical types of speakers with the expectation that one's gender

identity remains relatively fixed across interactions, unchanged from the gender ascribed by the sex assignment at birth (Hall & Davis 2021).

In post-structuralist models of gender identity (Baxter 2003; Holmes 2014), however, gender is recognized as fluid and socially constructed (Hall & Bucholtz 2012). Individuals' gender identity is not a static label, but rather an ongoing project that individuals themselves reify or disrupt through everyday interaction, expression, and practice (Butler 2013; Wodak & Benke 2017). Individuals enact understandings of their own and others' gender by presenting and performing social acts that are ideologically related to larger conceptualizations of gender (Butler 1990; Zimman 2012). As an example, Calder's (2019) work on San Francisco drag queen communities exemplifies the fluidity and performativity of gender: queer speakers, who most often were assigned as male at birth, recruited performative embodied resources – such as wigs, shapeware, and makeup – as well as particular sociophonetic features – such as /s/-fronting – that are aligned with their feminine drag queen personae. In this way, these speakers' sociolinguistic performances of the hyper-saturated, "larger-than-life" feminine drag personae challenge the male-female dichotomy and the gender roles it assumes of speakers assigned into male and female categories.

Furthermore, critical models of gender challenge a normative binary construction by providing space for exploring the linguistic practice of individuals whose gender identities lie beyond the binary categories of male and female. Zimman (2019) emphasizes that non-binary and trans communities have historically been erased and marginalized from social scientific research, sociolinguistic endeavors being no exception until relatively recent pursuits in the field (e.g. Kirtley 2015; Gratton 2016; Zimman 2018; Steele 2019; Calder & Steele 2019; Clifford 2019; Hazenberg 2019; Parnell-Mooney 2019; Zimman 2019; Eckert & Podesva 2021).

Emerging sociolinguistic studies have shown that non-binary speakers' pitch range and phonetic realizations do not pattern according to a binary model of what is predicted for male or female voices (Calder & Steele 2019; Scmid & Bradley 2019). Additionally, transmasculine speakers, or individuals in or post-transition to a male gender identity, in Zimman's (2017) study were found to recruit linguistic variables associated with binary gender categories in order to adhere to or subvert normative expectations of male versus female categories.

Meaningful differences among speakers within male and female gendered groups have also historically been ignored in research, furthermore essentializing a male-female dichotomy. Work of important exception that delves into these differences is Eckert's (1989) *Jocks and Burnouts* ethnography, which illustrated that variation in suburb-Detroit adolescent girls' productions of Northern Cities Vowel Shift features was guided by their involvement in different social groups. Stecker & D'Onofrio (2022) found that listeners exhibited greater variability in social evaluations within gendered categories of speakers than in opposition between groups of speakers perceived as male or female. In this way, the production, perception, and interpretation of gendered sociolinguistic performances is crucially mediated by social understandings of both the speaker and the listener, which complicates notions of fixed gendered categories.

Yet, essentialist notions of gender that naturalize a binary between male and female categories remain pervasive across all kinds of domains of human activity. Historically, both popular and academic spaces have focused exclusively on the similarity or differences between (rather than variation within or beyond) male versus female groups of speakers, perpetuating the social reasoning of an essentialized and cisnormative binary conception of gender (Meyerhoff & Ehlrich 2019). In particular, while Lakoff's (1973) work on was one of the first to provide academic space for describing the relationship between language and gender, her observations

and analyses were limited to a comparison exclusively between men and women, and much sociolinguistic research since has followed suit.

Listeners have been shown to perceive an individual's self-identifying gender along binary lines with relative perceptual ease and accuracy (Owren, Berkowitz, & Bachorowski 2007). For example, listeners in Mullenix's (2005) study replicate previous findings from Barreda (2017), Smith et al. (2007), and Smith & Patterson (2005) by more readily associating fundamental frequency information with vocal tract size and overall speaker size, which then is ideologically mapped onto binary gendered categories of speakers (men normatively occupying bodies of larger sizes and/or with longer and larger vocal tracts than women).

Biases reflecting static, binary conceptions of gender also occur in more automatic or implicit processes: In their social psychological work, Phillips & Boroditsky (2003) found that Spanish and German speakers, who have opposing grammatical gender assignments to nouns like *key* or *bridge*, illustrate gendered associations with those nouns with respective gender grammatical structure. For example, Spanish speakers, when presented with a masculine gendered noun such as *puente* (*bridge*), more frequently classified this item in the same category represented by a male-figure depiction rather than in a category represented by a female-figure depiction. Contrastively, German speakers, who had a feminine grammatical gender assignment for these items, assigned these nouns to the female-figure category. This research suggests that individuals' conceptualizations of objects and nouns may be biased to uphold gendered mappings inscribed in the grammatical gender systems of their native language.

Other sociolinguistic work illustrates binary-based gendered expectations for men's versus women's voices in more conscious metalinguistic practice. In one example, Slobe's (2018) study of "Mock White Girl" online performances explores moments in which girls'

voices are a site for parody. In her analysis, Slobe (2018) found that particular linguistic features were meaningful elements for young male performers on YouTube to imitate young, affluent, white, female speakers. Specifically, linguistic features such as creaky voice, uptalk, and *like* as a discourse marker were manipulated here based on their stereotypical links with women's voices (Lewandowski 2012; Chao & Bursten 2016). These virtual satirical performances focused solely on white women caricatures, erasing conceptualizations of other types of personhood co-occurring with this linguistic style. These parodic instantiations illustrate the ways in which speakers semiotically link attributes such as "unprofessionalism," "naivety," and "stupidity" to the vocal qualities of female voices, with the purpose of commenting on a caricaturized notion of femininity and 'white feminism', or feminist expressions that fail to address forms of oppression faced by non-white women (Ortega 2006). In doing so, these online speakers constructed the 'Mock White Girl' as a stylized personification of these attributes. These male-identifying Youtube performers reaffirm a binary lens of gender by explicitly switching from their 'Mock White Girl' performative voices to their own "authentic" styles as male-identifying individuals (Slobe 2018). The meta-discursive labelling of a 'Mock White Girl' style sets apart this specific demographic identity from all others while ideological attitudes about these voices simultaneously ascribe particular linguistic features to satirical stylizations in these performances.

Metalinguistic commentary described by Gross (2015) among other examples (Slobe 2016; 2018) illustrates the greater degree of scrutiny that women face compared with men when engaged in socially salient or stigmatized practices, from overt write-in or on-air criticisms that women on an American culture podcast received from their listeners in contrast to their male colleagues (Morgan 2015), to the online blacklisting of female Arabic celebrities for violating

expectations of what is “appropriate” in their professional linguistic landscape (Hachimi 2016). These examples reflect patterns of robust social expectations that women’s voices adhere to standards regulated not by themselves but by others – an artifact of patriarchal social norms that subordinate women’s gender roles in social spaces to objectification (Bartak 2015; Swami & Voacek 2013; Brunner 2013). These social norms leave little room for variation in expressions of female gender identity by emphasizing roles enforced by localized traditions and standards, furthermore relegating individuals’ gender expression and identities to performances expected by roles patriarchally associated with male and female categories.

In these ways, perceptions of what speakers of female and male genders “should sound like” derive from ideological expectations from these voices and mappings to these gendered categories. Though gender identity emerges through ongoing and interactive social praxis (Eckert & Podesva 2021), abstract gender categories are nonetheless in operation in implicit and explicit perception (Strand & Johnson 1996; Calhoun et al. 2023). These gender categories have predominantly been constructed through a cisnormative lens. In particular, categories of male and female genders are continuously constructed through ideological assignment of social values to these categories, until these categories are perceived as fixed and static to an individual (Butler 1990). The persistent circulation of comparisons between male and female speakers essentializes a binary distinction between these categories, rendering this conception of gender insidious in processes of socialization (Ochs 1992), and metalinguistic discourse on which gendered groups’ linguistic practice is worth noting and why (Nakamura 2014; Bridges 2017). These processes furthermore erase critical non-binary understandings of gender identity by ignoring gender identities other than those captured in a male-female dichotomy. Metalinguistic expectations of these gendered groups in US contexts predominantly derive from standards for voices perceived

as white, thereby casting voices racialized as non-white as deviant outliers to existing ideological structures (Rosa 2016). In these ways, cis- and white-normativity have been found to be pervasive in individuals' practices of interpreting linguistic gendered performances and expressions of their identity.

I call on male and female gender categories in this dissertation to inspect the ways in which acoustic cues to these social groups activate expectations about their use of a sociolinguistic feature, modulating listeners' speech recognition when listening to speakers perceived as women versus men. For clarity and consistency, I use "male" and "female" modifiers to refer to individuals (both speakers and listeners in memory experiments) who identify as men and women, respectively. As these terms denote sexed categories of speakers, I acknowledge that these terms for sex signify a different distinction than that between the terms "man" and "woman" for gender identity categories. Furthermore, while listeners in the memory experiments described below received no top-down category information on the gender identity of their speaker, categories of "male" and "female" genders entail listeners' perceptual classification of speakers into these groupings. As described below, speakers serving as experimental stimuli for listening participants were nearly categorically perceived as their self-identified gender. For this reason, I use (perceived) speaker gender as a term to describe speakers identifying as male or female. I acknowledge that drawing on stimuli from speakers described as "male" versus "female" signifies a binary distinction that does not reflect how gender identity is realistically constructed in social spaces. I use these terms for perceived speaker gender as they not only most closely align with the gender and sexed identities of speakers I recruited for stimuli construction, but also because they mimic the predominant ideologies related to language and gender that circulate in the popular imagination. In this way, the main studies of this

dissertation focused on the male and female speaker gender categories in order to examine the distinction between these genders that remains pervasive particularly in metalinguistic discourse about uptalk, described below.

1.4.2. Uptalk: a psychosocial linguistic phenomenon with gendered indexical values

Gender ideologies not only regulate perceptions and evaluations of speakers' sociolinguistic practice, but also the indexical field (Eckert 2008) of the social meanings that come to be linked with linguistic variants (Warren 2016, Tyler 2015, Calder 2019). As previously discussed, a listener's perception and interpretation of speech is fundamental to the ways in which linguistic forms are imbued with socially meaningful semiotic structures (Levon 2008; Eckert 2012). In particular, uptalk in American English has discursive functions including turn-taking negotiation, floor-holding, and requests for listener confirmation (McLemore 1991; Warren 2016). This intonational feature has been linguistically characterized as constituting variants of different contours, and relatively difficult to quantify and measure in linguistic research (Di Gioacchino & Jessop 2010). Intonational features, such as uptalk, have been defined as "the use of suprasegmental phonetic features to convey 'postlexical' or sentence-level pragmatic meanings in a linguistically structured way" (Ladd 1996:6). For example, uptalk as a rising intonational contour recruits the speaker's command of pitch accents and rising tones across the space and timing of linguistic utterances. The intonational contour in an uptalk production, however, can be instantiated differently depending on the physiological and social constraints of a speaker's voice and linguistic repertoire (Shokeir 2008; Asch & Brogan 2022). Uptalk has been shown to encompass different rising contours, codified by the Tone and Break Index (ToBI) system of annotation (Beckman, Hirschberg, & Shattuck-Huffnagel 2006): While some instances of uptalk can be represented by H*H-H% (contour with a high nuclear pitch accent and boundary tone, i.e.

relatively high F0 in the final stressed syllable), others can be represented by L*+H L-H% (contour with a rise at the terminal of the utterance as well as within the duration of the nuclear pitch accent In the utterance). Despite these empirical difficulties and variability of uptalk as a linguistic form, listeners can perceptually readily recognize uptalk as a specific form of rising pitch across different speakers, utterances, and specific realizations of the feature (Di Gioacchino & Jessop 2010).

Uptalk also bears stereotypical associations with young women (Gorman 1993; Marsh 2006; Horowitz 2006; Duam 2007), often described as "sorority speech" or "Valley Girl speech," explicitly linking this feature to "ditzy-ness" or the linguistic styles of "rich, white young females from the San Fernando Valley" (Ritchart & Arvaniti 2014; D'Onofrio 2015). These associations between greater frequencies of uptalk use and female speakers occur both in metalinguistic commentary (Slobe 2018) and in linguistic experimental work (Tyler 2015). For example, Calhoun, Warren, Mills, & Agnew's (2023) study tested New Zealand listeners' mappings of uptalk tokens and male versus female gender in an Implicit Associations Test (IAT). Listeners were more likely to associate (digitally manipulated) uptalk tokens with female-gendered names than male-gendered names, and this gender effect was modulated by the perceived gender of the speaker, listeners' own gender identities, and their beliefs and attitudes about gender. Specifically, female-identifying participants exhibited a stronger associations, between uptalk and female names than male-identifying participants, indicated by difference in their reactions times between stereotypically congruent (e.g. uptalk token from female speaker) and incongruent trials (uptalk token from a male speaker) in the IAT. Further, female-identifying participants with attitudes against endorsing Benevolent Sexism (Glick & Fiske 1996) had longer reaction times in incongruent trials than in congruent trials, in contrast to male-identifying

participants and female-identifying participants with attitudes biased toward Benevolent Sexism. In this way, individuals' existing interactions with gender – their own identity and the stances they put forth – modulated how they processed tokens of uptalk.

Other recent sociolinguistic work in production provides evidence speakers of all genders use uptalk (Linneman 2013). This process is also mediated by social pressures that condition women to linguistically adopt uptalk and other gendered features to interact in their linguistic marketplace (Bourdieu 1977; Eckert 1989a), leading to greater rates of production by women in various social spheres (Warren 2016). In particular, the linguistic socialization that girls experience exhibits a prevalent theme in teaching these individuals to navigate "appropriate" articulations of desire and interpersonal conflict, preparing them as social actors to enter the "heterosexual marketplace," in which the social indexicality of language serves social mobility via linguistic articulations of sexuality, social attractiveness, and desire (Eckert 2003, 2011, 2014; Bucholtz 2009; Mendoza-Denton 2014). As an example, Linneman's (2013) analysis of uptalk use among *Jeopardy!* contestants found that when male participants were leading the round, they used uptalk at lower rates than female participants at points in their game leads. He reasoned that while women use uptalk at greater rates when winning as a way to hedge their own success in their gaming context, male speakers avoid this feature when winning to reinforce masculine roles of competitiveness. In this way, previous work illustrates that uptalk is conversationally operationalized depending on listeners' existing self-conceptualizations and desires for how to be perceived in an interactional context.

1.4.3. Ideological influences in the indexical field of uptalk

The social meanings in the indexical field (Eckert 2008) of uptalk, like other sociolinguistic features, emerge from continuous interactional processes in the space between speakers' practice

and listeners' interpretations (Podesva 2007; Eckert 2012). Listeners make evaluations of gendered voices at different levels of awareness (Eckert & McConnell-Ginet 2003). While metalinguistic commentary hosts a multitude of explicit attitudes about sociolinguistic performances of gender, gendered biases may also be less overt or direct. For example, when interviewed by the *New York Times*, linguist Cameron Fought highlighted distinct patterns in how societal audiences interpret women's use of stigmatized features: "If women do something like uptalk or vocal fry, it's immediately interpreted as insecure, emotional or even stupid" (Quenqua 2012, online).

In interactive evaluations and stereotypical discourse about uptalk, speakers and listeners circulate indexical links that a) naturalize women's use of uptalk as essential to their gendered social type and b) erase (Irvine & Gal 2000) notions of gender identities beyond a binary cisgender categorization. The indexical structure of uptalk circulates cyclical links between uptalk and femininity: speakers who adopt uptalk into their linguistic styles are perceived with the same qualities ideologically ascribed to femininity, such as "friendliness" or "insecurity." For example, a gay male speaker in Podesva's (2011) study style-shifted with his use of uptalk to mediate the perceptual activation of his "caring-doctor" persona. These results suggest that interpretations of uptalk are confined to locally situated understandings of the context, including the relationships between interlocutors and specificities in how listeners ascribe gendered social types onto their speaker.

At the same time, listeners' perceptions that a particular speaker belongs to the female gender category guide the ways in which components of the speaker's linguistic style, including their use of uptalk or even fine-grained phonetic distinctions, is in turn assigned a feminine interpretation (Strand & Johnson 1999). It is then possible that uptalk has garnered so much

metalinguistic attention and salience in part due to its stereotypical linkage to women's voices and the greater degree of societal scrutiny of women's sociolinguistic practice in contrast to men (Morgan 2015; Quenqua 2012). Ainsworth (1993) states that the interpretation of upalk as denoting incredulity or uncertainty is a "gender-linked paralinguistic trait" (p. 282), with the rationalized connection between uptalk and powerlessness "intimately connected to the subordinate position of women" (p. 284). At the same time, uptalk has also been linked with positive traits in women's speech styles, such as its connotations to "the facilitative patterns which characterize women's speech in general" (Holmes 1993:111).

This same ideological guidance and social expectations of speakers mediate uptalk's negative and stereotypical links to "ditzy", "vapid", or "uncertain" attributes (Ainsworth 1993; Guy & Vonwiller 1984; Tyler 2015). Through this social reasoning, features of women's speech are expected to convey those qualities assigned to them via existing patriarchal social structures that center men's voices as a default – a kind of invisible norm while positioning women's voices as marked and as calling for linguistic policing and self-monitoring (O'Grady 2005). Examples of lived experiences from female-identifying individuals illustrate that women undergo practices of silence – the absence of participation in an interaction as a social actor – and self-surveillance of their own sociolinguistic performances to navigate hegemonic power structures in their particular communities (Babel 2016).

As another example, Levon & Ye (2019) tested listeners' social perceptions of this feature and subsequent evaluations of their speaker by inducing a highly charged political and gendered context: a mock rape tribunal between a male defendant and female complainant. They found that listeners' interpretation of the male defendant's uptalk provided positive social attractiveness, while their interpretations of (the same rate of) uptalk use from the female

complainant did not enjoy these advantages. This provides evidence that processes in sociolinguistic perception, such as the noticing and interpretation of a speaker's use of uptalk, is guided not only by the context in which speech occurs, but also by "the gender ideologies [that] constrain who has the authority to speak, to be heard and to be believed in different contexts" (Levon & Ye 2019:16).

For English listeners in the US, existing social expectations about who uses uptalk are products of this ideological frame that confines metalinguistic commentary about this feature to a binary lens of gender. For example, ongoing metalinguistic discourse has typified uptalk as an iconic trait of (cisgender) women's voices, exemplified in the codification of the 'Valley Girl' persona (Ritchart & Arvnaniti 2013; Nycum 2018) and metadiscursive definitions of uptalk as "a part of women's nature" (Green 2014, online). In this way, listeners' understandings of uptalk are often grounded by historically ideological conceptions of cisnormative (marginalizing input from trans-speakers and non-binary speakers [Zimman 2017; Gratton 2016]) gender as a dichotomous classification that is permanent and unchanging through one's lifespan (Snyder et al. 1977).

While speakers of all genders use uptalk, commentary in meta-discursive spaces pervasively compare the use of this feature between binary male and female gendered groups (Gross 2015). Commentary in online forums and pre-digital spaces have historically disparaged uptalk as a vocal tic requiring rehabilitation, often directly targeting women for speech therapy to eliminate the use of uptalk (Masterclass 2021). Some instantiations of commentary have cited women's presupposed greater usage of uptalk in contrast to men as reasoning for professional and income inequality (Anderson et al. 2014). Within the last decade, rebuttals against the criticism of uptalk have emerged in metalinguistic discussion, arguing that the scrutiny against

uptalk is a patriarchal artifact of denigrating women's voices and features associated with them (e.g. Seitz-Brown 2014; Gross 2015; Luu 2017). Nevertheless, these discussions by and large circulate a comparison between speakers of two gendered categories: those described as "men" or "women" (e.g. Thomson 2018; Green 2014; Cameron 2015).

The goal of this dissertation is to delve into the gendered biases individuals have about uptalk, manifested in the popular discourses of this intonational contour as a gendered linguistic feature. Given popular discourses surrounding the use of uptalk linked with women's speech styles (Green 2014; Dziura 2020; Baldoni 2015; Davis 2010), this dissertation thus attends to a binary comparison between speakers heard overwhelmingly as white women or white men. I acknowledge this choice limits the analyses from incorporating speakers of other racialized and gendered groups. While studies in this dissertation examine the social expectations predominant in mainstream discourse, this choice in focus may perpetuate a binary conceptualization of gender. Conclusions from these studies encourage further empirical pursuits that delve into perceptions and speech recognition of non-binary speakers, the intersection of gender identities with sexuality and racialized identities, along with speakers grouped by other socially meaningful dimensions.

1.5. Variation in listener attitudes and expectations

Individual listeners exhibit variation in their ideological frames about gender, further undermining the static notions of a male-female gender binary. Listeners' expectations and attitudes are largely shaped by their personal lived experiences, as well as through less-direct nonpersonal methods: Social stereotypes in media (Hargreaves & Tiggemann, 2003; Levina, Waldo, & Fitzgerald, 2000), one's inherited family values (Olson, Vernon, Harris, & Jang,

2001), and individuals' interconnectedness to others' lived experiences (Albarracin & Shavitt 2018) all complicate the construction and maintenance of individuals' attitudinal and social lenses. In this way, an individual's own positionality mediates their interpretations and attitudes in their social world.

1.5.1. Positionality of the researcher and listeners

The term *positionality*, used heavily in cultural and linguistic anthropology, has typically been used to describe the predominance and influence of individual academic researchers' own backgrounds and ideological biases that manifest in their own work (Harding 1987; Jacobs-Huey 2002; Okely 2012; Holden 2022). In this dissertation, my own positionality as a researcher is an important factor in every step and methodological decision in these studies. My research focus on gender and constructs in linguistic memory is borne from my own lived experiences of identifying as and being perceived as a woman. The selection and construction of stimuli were limited to using resources immediately available to me as a graduate student living in Chicago during the COVID-19 global pandemic. The analyses and interpretation of data collected in these studies are shaped by my education that cultivates deep interest in the social structures underlying ongoing linguistic variation and mental representations of language (Labov 1974; Eckert 2012; D'Onofrio 2021).

In these ways (among others), as much as this dissertation is an analysis of others' linguistic performances, it is also a product of my own positionality and could never exist in an objective space immune to my immediate placement in my social and academic worlds, and the prominent forces (like social isolation from a global pandemic) within them. I have had the experiences of identifying with the female-sexed category of "girl" that I was assigned at birth and at times have been overtly cognizant of the educational, familial, and popular discursive

forces that created the category of "woman" into which I was socialized. As a cisgender white woman in particular, white-supremacist and cisnormative ideological practices in the US sociolinguistic, political, and economic landscape (Bucholtz 2019; Rickford & King 2016; Calder 2020) do not target, police, or marginalize my physical body or linguistic style in the same way that they do other individuals with different positionalities. Rather, I have been afforded opportunities as a cisgender, white, able-bodied person to receive a position in a doctoral program at a leading research university. At the same time, patriarchal forces in my emergent socialization into feminine personhood, especially as a late millennial (Hess 2013; Wolf 2015; Nunberg 2016), have often yielded overbearing and unsolicited attitudes and expectations toward my voice (Chao & Bursten 2020), while also protecting my male-identifying peers from the same levels of scrutiny (Anderson & Klofstad 2012). It is through these opportunities and consequences that I've experienced as this type of person that provide me exclusively with my unique positionality – my encompassing experiences, socialization, and training – with which I collect and interpret memory and perception data from listeners.

Here, I operationalize positionality as the way in which individuals are socially oriented in relation to their surroundings, including other social forms such as social types, institutions, and stereotypes, and the influence of this framing as they map to and interact with their social worlds (Elie 2006; Bialecki 2021). The existing experiences, beliefs, and attitudes that an individual uses to locate and identify themselves with others are meaningful components of their sociolinguistic background, and they continuously operate as a positional frame.

As an example, Niedzielski (1999) illustrated how Michigan speakers' expectations of their own accent versus a Canadian accent conditioned their perception of raised and un-raised diphthongs. Components of Michiganders' positionalities – their experiences with and

understandings of these sociolinguistic features in their social landscape – guided these cognitive processes. Linguistic research needs to thoroughly understand the power of listeners' social positionings not just in their speech perception performance but also in their access to and interaction with fundamental cognitive understandings of where they are in their linguistic landscape (who is talking, where, and when). This dissertation explores specific components of individuals' positionality around uptalk, namely their personal experiences and attitudes around this feature, and it further investigates whether these experiences and attitudes are influential in listeners' recognition of falling and rising utterances from male and female speakers.

While my dissertation investigates listeners' recognition of speech produced by speakers who self-identify and are perceived as men and women, this project also uses quantitative survey paradigms to explore individual listeners' evaluations of their speaker within these gendered groups, providing more dimensions for different gendered expressions and interpretations for analysis. Studies 1 through 3 in this dissertation first use a forced-choice social evaluation survey: After listening to each speaker in their sample, participants indicated how they evaluated each speaker on a variety of different social attributes, along with their perceptions of the speaker's demographic qualities. Qualities for speaker evaluation questions derived from existing indexical links with uptalk, such as social attractiveness or popularity, sensitivity to listener engagement and comprehension (Levon & Ye in press; Tyler 2015; Warren 2016). Responses to these questions elucidate meaningful differences in how particular male- and female-perceived voices are socially assessed along these uptalk-related dimensions.

1.5.2. Quantification of listener positionality

This dissertation utilizes another survey after the memory tasks as an opportunity to delve into the sociolinguistic backgrounds of listeners: After listening to their set of speakers and

completing the memory tasks, participants responded to a series of introspective questions. This sociolinguistic background survey enabled me to delve into individual listeners' positionalities regarding their own personal lived experiences with uptalk as well as their expectations toward others' voices, such as the use of uptalk in public professions. The content in these statements highlight social values that share indexical links with uptalk. For example, Tyler's (2015) perception study showed that listeners associate uptalk with insecurity, lack of professionalism, and either the continuation or termination of one's turn[1]. Survey questions constructed in this dissertation direct listeners to describe their own sociolinguistic experiences and attitudes around the existing indexical field of uptalk.

The studies in this dissertation address how the dynamic and nuanced interpretations of gendered speakers can vary not only by the speakers themselves, but by what listeners are bringing into their evaluations and participation in the experiment. More specifically, I probe the variation of listeners' perceptions of their own voice as well their attitudes toward others' voices, and I in turn analyze these components of listeners' backgrounds in influencing their recognition of a stereotypical gender-linked linguistic feature, uptalk (specifically in Chapter 5). By using rising intonation interpreted as uptalk as an experimental variable in stimuli from male- and female-perceived speakers in a series of memory tasks, these studies provide a window into the goings-on of the cognitive workings behind sociolinguistic perception. In this way, this dissertation contributes to a more thorough understanding of the relationship between ideological constructs such as gender and the linguistic mental representations emergent in sociolinguistic production and perception.

[1] The paradigm in the memory experiments in Studies 1 through 3 (Chapters 2 through 4) control for the continuation of a speaker's turn by confining the listener's context to a single speaker's podcast recording, rather than a multi-interlocuter turn-taking context with more than one speaker sharing the conversational floor.

An examination of the complex and dynamic intersection of gender ideologies and sociolinguistic cognition is particularly compelling because, as shown in metalinguistic commentary (Slobe 2016, 2018; Hachimi 2016; Gross 2015; Morgan 2015), speakers and listeners attend in great detail to sociolinguistic performances of gender and scrutinize such performances in accordance with their gendered expectations. If social ideologies and the expectations derived from them impact the degree to which listeners attend to sociolinguistic features, then I hypothesize that the perceived gender of a speaker and the social expectations about who is likely to use uptalk have great potential to influence listeners' memory of this feature. In turn, this may then affect how listeners cognitively store sociolinguistic input from speakers and provide evidence for ideologically biased influences on linguistic memory.

1.6. Organization of chapters

In three experiments, I draw upon gender ideologies and a gendered sociolinguistic feature, uptalk, to investigate whether social expectations and gendered stereotypes impact listeners' memory of sociolinguistic information. I use an old-new judgment task in which listeners classified whether specific utterances constituted the same or different intonational contour than heard previously in a training passage. I first assess whether listeners' old-new recognition of rising utterances perceived as uptalk is different between speakers perceived as either men or women in Study 1 (Chapter 2).

To more closely examine how these cognitive biases may be evident in listeners' memory, Study 2 (Chapter 3) assesses the degree to which listeners exhibit false memories of rises interpreted as uptalk when their speaker never produced any rising intonational contours. Rates of false recall, or instances in which listeners inaccurately recognize some linguistic form

as “old,” have proven to be a meaningful site for previous empirical linguistic memory work. While D’Onofrio (2021) found these false alarm effects for a phonetic feature linked to a geographic social category of speakers (Californians) and a social persona (Business Professional), I aim to replicate effects of listeners’ social expectations on their recognition of a gendered linguistic stereotypical feature when listening to speakers perceived as men versus women. While Study 1 investigates the degree to which the mapping of women’s voices and uptalk garner greater rates of false alarms for women’s rising tokens in contrast to men’s rising tokens, Study 2 examines baseline thresholds for falsely recognizing rises when none were produced by speakers perceived as men or women.

Furthermore, Study 3 (Chapter 4) assesses the influence of metalinguistic commentary by employing an experimental prime that induces explicit ideological expectations about particular speakers’ voices. As previous work (McGowan & Babel 2019) points to complex influences of expectations prevalent in metalinguistic discourse on listeners’ memory of speech, further clarification is needed to assess whether a metalinguistic attitude about a stereotyped linguistic feature can be experimentally primed, and can induce a bias on listeners’ recognition of different speakers’ rising utterances. Study 3 utilizes priming statements comprising metalinguistic commentary about gendered expectations of which social groups of speakers use uptalk. Priming listeners with information about which speakers are likely to use a gendered sociolinguistic feature may influence how they attend to and recognize those productions from speakers perceived as either male or female. Analyses for Study 3 compare the effects of different metalinguistic primes to the absence of any metalinguistic information listeners’ performance in (false) recognition of utterances.

In Chapter 5, I delve into listeners' own varied sociolinguistic experiences and the degree to which their existing attitudes endorse standard language ideologies about others' voices and uptalk in public and professional spaces. The analyses in Chapter 5 further incorporate these aspects from listeners' existing sociolinguistic background as factors in how listeners' positionality maps expectations about gendered voices onto incoming speech streams and linguistic cognitive representations. Specifically, Chapter 5 analyzes the relationship between factors in listeners' own sociolinguistic backgrounds and their recognition behavior in the old-new recognition task. Finally, in Chapter 6, I discuss the implications of my results for theories regarding the relationship between social meaning and linguistic forms in cognitive representations.

2. Study 1

2.1. Introduction

Study 1 investigated how listeners' memories of linguistic productions are impacted by the social meaning embedded in the linguistic signal and their perceptions of their speaker. Previous theoretical and experimental work illustrates the relationship between social factors and processes in which speech is stored and represented in the mind (e.g. Strand 1999; Niedzielski 1999; Hay & Drager 2010; Walker et al. 2019). Specifically, listeners have been shown to retain information regarding both their particular interlocuter as well as their linguistic productions in cognitive representations (Bradlow et al. 1999; Church & Schacter 1994; Goldinger 1996; Nygaard et al. 1995; Schacter & Church 1992). Sumner et al.'s (2014) work theorizes that existing social knowledge may strengthen the encoding of incoming linguistic information that is linked with particular social concepts in cognitive representations, leading to a stronger, more robust representation of these linguistic forms in speech perception. However, less is known about what these social concepts are, and how they influence listeners' memory and cognitive representations of linguistic forms. Further investigations of the ways in which listeners recognize sociolinguistic forms from explicit memory are required to understand the link between speech and embedded social meaning in cognitive representations.

Study 1 tested listeners' recognition of rising versus falling utterances, both socially meaningful prosodic contours, when listening to speakers of different perceived genders. In this study, rising utterances served as acoustic instantiations of uptalk, a psychosocial phenomenon. That is, uptalk can be realized by several intonational contours (Prechtel & Clopper 2016), and various components of a listener's sociolinguistic background and perceptions of a speaker have been shown to influence how they notice and interpret a speaker's use of uptalk (Di Gioacchino

& Jessop 2010). This dissertation used stimuli comprising declarative sentences produced with a rising contour and high final boundary tone as a set of utterances that were predominantly perceived as containing uptalk. In this study, participants listened to a set of speakers (two perceived as male, two perceived as female) produce an informational reading passage. Then, they performed two tasks: first, participants completed an old-new recognition task, in which they heard a series of single utterances and indicated whether each utterance was "old", in the sense that they previously heard that recording in the passage, or whether the utterance was "new", or a new version/recording of the utterance.

Old-new judgment tasks have been employed in investigations of memory performance with spoken and written language (e.g. Fine, Jaeger, Farmer, & Qian 2012; Zormpa et al. 2019). As an example, Craik & Kirsner's (1974) study investigated continuous recognition for different spoken words and voices. They found that listeners' word recognition improved when trials were consistently produced with the same versus different voices. In their (1993) modulation of this study, Palmeri, Goldinger, and Pisoni's work illustrated that listeners retain social information about their speaker embedded in their voice to aid in recall and recognition of linguistic forms. These experimental insights suggest that listeners store speaker-specific acoustic cues in their mental representation of language, motivating further inquiry in the composition of socially meaningful linguistic forms in memory.

The old-new recognition paradigm has also been used to examine the ways in which extralinguistic factors, such as the social meaning imbued in a linguistic form and the frequency with which the form is produced, affect recognition. In their (2005) experiment using the old-new judgment paradigm, Sumner & Samuel tested listeners' recognition of lexical items with different word-final [t] variants. Listeners' old-new judgments illustrated that the "canonical"

aspirated variant of word-final [t] exhibited greater long-term memory effects in contrast to other more frequently produced [t] variants (glottalized or coarticulated variants). In this way, listeners appear to be sensitive to the various social associations of fine-grained phonetic details in the ways in which they retain these linguistic forms in memory. Other sociolinguistic experiments have also modulated this old-new judgment paradigm to manipulate top-down information about the speaker (e.g. Niedzielski 1999; Hay et al. 2006; D'Onofrio 2016, 2021) or bottom-up information encoded in the acoustic information without any top-down information (e.g. Giles 1970; Goldinger 1996; Sumner & Samuel 2005) to test listeners' perception and recognition of linguistic forms. For example, D'Onofrio's (2021) work found that listeners' recognition of tokens of the backed-TRAP vowel, a feature linked with supra-local geographic orientation (Podesva et al. 2012) and to the "Business Professional" persona (D'Onofrio 2015), was greater when a speaker was labeled as a "Business Professional" in contrast to those provided with no speaker information.

Previous experimental work has therefore provided grounds to explore additional speaker- and listener-specific factors in linguistic memory. Unlike previous studies, Study 1 employed this old-new judgment paradigm to test listeners' recognition for intonational contours, rather than to investigate acoustic factors at the segmental level in listeners' lexical access (Goldinger 1996; Sumner & Samuel 2005). While some recent experimental work has elucidated the perceptual side of uptalk (e.g. Tyler 2015; Tobin & Benders 2018; Wollum 2019; Calhoun et al. 2023), the functions, acoustic forms, and meanings of this intonational tune have predominantly been established through studies of production by different speakers and/or in different contexts (e.g. Shokeir 2008; Di Gioacchino & Jessop 2010; Ritchart & Arvaniti 2014; Prechtel & Clopper 2016; Asch & Brogan 2022; Song, Clopper & Wagner 2022). The old-new

recognition task in Study 1 is the first to my knowledge to implement a memory task for listeners' recognition performance of rising contours previously perceived as uptalk.

In the second task of Study 1, participants completed a cloze (word recall) task in which they recalled what the speaker said by filling in missing information from the passage they just heard. Previous studies using this paradigm have investigated listeners' speech recall with regards to different language background factors, such as listeners' experience with a language variety and the speaker's perceived racialized identity (Rubin 1992). Study 1 implemented this task to examine whether listeners exhibited different accuracy rates of recalling content when listening to speakers of different perceived genders.

After the memory tasks, participants completed social evaluations for each speaker and finally answered some questions about their own sociolinguistic attitudes and background. Participants' responses across this study's components were submitted to a series of analyses to identify significant predictors in listeners' memory of falling and rising contours produced by speakers heard as one of two different genders: men or women. Results illustrated that previously heard women's rises were most frequently recognized in contrast to men's rises and women's falls, suggesting a gendered bias in memory of intonational contours that maps to broader metalinguistic stereotypes about the gendered patterning of uptalk.

2.2. Stimuli

2.2.1. Passage

Three passages were written as potential stimuli for the main experiment. Each passage focused on trivial non-fictional information about a commonplace topic (frogs, keychains, jigsaw puzzles).

In a preliminary normalization study Norming Study 1, 20 participants[2] read each passage and then evaluated the passage on a series of scales: Participants were asked how emotionally charged and entertaining they perceived the story to be, and how likely it was that they could hear this story told by a male, female, or non-binary speaker, each evaluated on a scale from values of 1 (not at all charged/entertaining/likely) to 5 (very charged/entertaining/likely). These characteristics were selected in order to establish whether the content would garner more relative attention due to listeners' sensitivity to any political or gendered associations with the passage. Results from this preliminary task showed that listeners interpreted the frog passage to be least biased in the terms of being heard by a speaker of any particular gender and least emotionally charged, so this passage was chosen as the stimulus for the main experiment. Table 2.1. below reports the average ratings and standard deviations for evaluations for each passage. The full passage is included in Appendix A.

Passage	Emotional charge mean	Entertaining Mean	Male bias Mean	Female bias Mean	Non-binary bias Mean
Frog	2.870	4.130	2.652	3.783	3.435
Jigsaw	3.043	4.000	2.826	4.087	3.870
Keychain	3.652	4.130	3.783	4.130	3.478
Passage	Emotional charge St. Dev.	Entertaining St. Dev.	Male bias St. Dev.	Female bias St. Dev.	Non-binary bias St. Dev.
Frog	1.486	1.817	1.335	1.783	1.903
Jigsaw	1.665	1.931	1.435	2.065	1.74
Keychain	1.774	1.817	1.999	1.766	1.504

Table 2.1. Mean responses and standard deviations for each evaluation for every passage tested in the Norming Study 1.

[2] All participants for norming and main studies were recruited through Prolific, an online crowdsourcing platform for workers to complete tasks from their remote location – for more information, see Section 2.5.

2.2.2. Speakers

2.2.2.1. Recording sessions

A sample of 10 speakers was collected through convenience sampling of my social and academic networks in Chicago. Each speaker was recorded with the researcher to produce stimuli in a private, quiet remote location with an Audio-Technica PRO 70 microphone and Zoom H4n Pro hand recorder. Speakers were instructed to use a semi-professional style as if speaking on a podcast.

Speakers recorded four versions of the stimulus passage. In the first recording, speakers first listened to a model of each utterance provided by the researcher before repeating it. This was to control for the rate of speech and onset of the uptalk token for all speakers during each recording session.

In the first two recordings, speakers produced each utterance in a falling tone. After completing two falling tone recordings, each speaker first listened to the researcher provide a model for an uptalk utterance before recording the passage. The researcher continued to provide uptalk utterance models as needed until each speaker recorded two uptalk versions for each utterance in the passage. Each speaker paused for at least one second in between each utterance to aid in the subsequent concatenation of utterances in stimulus construction.

The modeling of the stimulus recordings by myself as the researcher delivers some opportunity costs. The researcher model for stimuli was implemented in order to control for the consistent patterning of several suprasegmental variables across the sample of speakers, such as the rate of speech, intensity, and onset of the rise or falling contour in each utterance. The researcher model prompted one instantiation of rising and falling productions for each utterance, which, in the case of rising stimuli, limited the analyses of this dissertation to listeners'

recognition of the specific intonational contours produced. This set of recordings for stimuli does not encompass all possible instantiations of contours that can be heard as uptalk. While additional norming studies described below sought to delimit the main task stimuli only to those perceived as uptalk tokens by listeners, these steps could not holistically guarantee that each utterance was consistently interpreted as an utterance containing uptalk. This limitation should be considered in future research.

Furthermore, speakers' productions could have adopted other suprasegmental or stylistic features present in the researcher's model, such as mean F0, or segmental features, such as mean center of gravity of sibilants. Speakers' unintended adoption of other features from the researcher's model may have led to unforeseen acoustic properties in speakers' productions of each utterance that could potentially have affected phonetic and/or social cues present in these utterances. However, while digital manipulation and construction of falling and rising utterances could have prevented these issues, the digital overlaying of a pitch contour could have also introduced potential complications in which manipulated productions may have sounded "unnatural" to listeners, influencing their social perceptions of speakers from their acoustic stimuli (Campbell-Kibler 2006) and, potentially, the memorability of certain utterances. In this way, the use of un-manipulated stimuli attempted to most closely emulate a feasible context for listeners to hear and recognize utterances with life-like rising and falling contours. Naturally produced utterances that emulated the researcher model were thus chosen, though the potential confounds introduced by this choice should be considered in interpreting study results.

2.2.2.2. Speaker selection

In Norming Study 2, 50 participants listened to all ten speakers' recordings of the stimulus passage. Each speakers' recording comprised nine utterances from the stimulus passage, with three utterances containing a rising contour.

After listening, participants evaluated each speaker along a series of social and demographic attributes. Specifically, participants indicated their perceptions of the speaker's gender, age, geographic region of origin within the US, racialized identity, and prototypicality of their voice as belonging to a man, woman, or non-binary speaker. They also indicated whether they perceived the speaker to sound like a native American English speaker, then rated each speaker for the following attributes: professionalism, friendliness, engagingness, attractiveness, nerdiness, intelligence, enthusiasm, and rate of speech (fast or slow). These attributes were selected following previous work on social perceptions of uptalk (Tyler 2015; Warren 2016)

Participants also indicated how perceivable each speaker's uptalk sounded. First, they received the following prompt: "*uptalk* is defined as the rising tone at the end of declarative sentences. When a speaker uses uptalk, it usually sounds like a yes/no question, even though the speaker isn't asking something, but rather making a statement. How perceivable was this speaker's uptalk?" Participants then indicated the extent to which they found the speaker's uptalk perceivable, scaling from 1 to 5 with labels ranging from "I didn't notice it at all" (1) "I noticed it once or twice" (3) and "I noticed immediately and constantly" (5). To measure whether listeners' classification of each speaker's rises as uptalk was consistent, listeners then heard one novel rising utterance and indicated whether they perceived the given clip as containing uptalk or a falling tone. Listeners' average responses for all social evaluations, demographic perceptions, and uptalk questions are included in Appendix B.

Based on results from Norming Study 2, six speakers' recordings were selected for stimuli creation in the main task. Rising utterances from one male and one female speaker (TS and AS) were accurately perceived as uptalk by less than 50% of norming study participants, so these two speakers were excluded. One male speaker (PS) was excluded as he was the only speaker who was perceived to sound more than 40 years old, and one female speaker (DD) was excluded due to significantly lower social evaluations relative to all other speakers.

According to Norming Study 2, the six remaining speakers were perceived to be 24-30 years old, white, non-U.S. Southern native American English speakers, whose rising productions were perceived as uptalk by at least 70% of participants in this preliminary norming study. Speakers were unanimously perceived by listeners as their self-identifying gender, with the exception of one speaker (JS) by one listener, who perceived her as a non-binary speaker. For this reason, I use the terms "male/female speaker" to describe these speakers for simplicity, as these descriptors apply to both self-identification and to Norming Study 2 perceptions. Listeners predominantly racialized all six speakers as white (over all other US census racialized categories), with perceptions of US regional origin split across the Midwest, West, and Northeast regions. Tables in Appendix B include perceived and self-identified demographic information for each of the six speakers.

2.2.3. Utterances

The stimulus passage that each speaker recorded comprised 53 sentences, 32 of which were target utterances used in the old-new recognition task (mean syllable count of target sentences = 12.9; SD = 3.94).

Each utterance served as individual intonational phrases. Of the 32 utterances in the stimuli, prosodic contours of 16 utterances ended in a falling tone, and 16 utterances ended with

a rise, authentically produced (unmanipulated) by each speaker. Each prosodic contour, falling or rising contour, occurred at the end of the intonational phrase with either a low or high boundary tone, respectively.

2.2.3.1. Utterance normalization

In Norming Study 3, each recorded utterance was judged by 40 naïve listeners as either containing "uptalk" or a "falling tone," given the following instructions: "For each recording, please indicate whether you heard the speaker use uptalk or a falling tone of voice. For this task, *uptalk* is defined as the rising tone at the end of declarative sentences. When a speaker uses uptalk, it usually sounds like a yes/no question, even though the speaker isn't asking something, but rather making a statement." Utterance contour (rise or fall) and speaker (x6) were counterbalanced across four conditions so that each utterance was heard and rated by 10 participants. Only utterances that were evaluated as uptalk or falling tone with 70% accuracy or greater were included as stimuli. Norming Study 3 listeners' average accuracy rates for each utterance's falling and rising versions are included in Appendix C. The average accuracy rate for each utterance was included as a factor in post-hoc analyses for each study, described in Chapters 2 through 4. In this way, the degree to which each rising utterance was perceived as uptalk and each falling utterance was perceived as a fall was considered as a main predictor in whether listeners classified that utterance as "old" or "new."

2.2.3.2. Acoustic analysis

As described above, I employed two types of intonational contours in declarative utterances to assess whether rising utterances were recognized more frequently when listening to speakers perceived as female versus male. Since each speaker's rising and falling utterances were authentically produced, an acoustic analysis was conducted on all speakers' utterances to

investigate any outliers and patterns in the ways in which speakers produced rises and falls in each utterance.

First, to analyze the characteristics of the intonational features in speakers' rising utterances, a set of pitch characteristics for each speaker's utterances was extracted via Praat. Specifically, mean F0, utterance onset F0, and F0 of the last stressed vowel for each utterance was recorded. The F0 measurement at the rise onset, characterized by the elbow of the rise trajectory showing a start of the clear upward trajectory of the contour (Levon 2020; Arvaniti & Ladd 2009) and the rise peak, characterized by the terminal boundary tone, were also extracted for each utterance. All F0 measurements were converted to equivalent rectangular bandwidth (ERB) to control for the auditory perception of F0 information.

Each utterance's rise excursion was also recorded to characterize the degree of pitch change and was calculated in two ways. First, speakers' absolute rise excursions for each utterance were calculated by the difference between the F0 at the rise onset and at the rise peak (Figure 2.1). Speakers' relative rise excursions were also calculated by measuring the absolute rise excursion in relation to the speaker's pitch range in each utterance (F0 max – F0 min). Finally, each utterance's rise slope for each speaker was also calculated by measuring the absolute rise excursion in proportion to the duration of the rise in the utterance (Figure 2.2).

Four separate linear regression models analyzed the effect of speaker on their relative rise excursions, absolute rise excursions, rise slope, and F0 range. Speaker coefficients were determined by comparing each speaker's deviation from the grand mean. Utterance was included as a random intercept. Full results from the analysis for these for F0 characteristics are included in Appendix D. Results from this acoustic analysis found some significant differences among speakers in the measured pitch characteristics. Speaker gender was also found to be a significant

predictor of speakers' F0 productions as well. Summarizing this data, speaker gender and individual speakers were found to be significant main predictors of nearly all F0 characteristics. First, male speakers were more likely to produce lower relative and absolute rise excursions than female speakers. Male speaker AV was found to produce rise excursions significantly lower than the grand mean than any other speaker. Male speakers were also more likely to produce rises with smaller slopes than rises produced by women. Female speakers ES and JO produced steeper rises in relation to the grand mean. Speakers' F0 ranges were not significantly predicted by speaker gender. Male speaker AV and female speaker ES were found to elicit larger F0 ranges in contrast to the grand mean, while male speaker GF and female speaker JS elicited smaller F0 ranges in contrast to the grand mean.

This acoustic analysis was not designed to find any exclusionary criteria for eliminating any of these six speakers from stimuli purposes. Rather, this analysis illuminated the variation of speakers' F0 space in their rising utterances. Different acoustic cues in rises have been found to be linked to variation in listeners' perceptions and processing of these utterances (Hirschberg & Ward 1992). Listeners have assigned a range of social values to rising contours, such as associations with "incredulity" and "uncertainty" (Warren 2014; Cruttenden 1981). These effects may manifest in variation with which rising and falling contours are recognized from memory. The memory effects of F0 information in stimuli utterances were examined in post-hoc analyses of listeners' old-new responses described in analyses sections for Studies 1 through 3 (Chapters 2 through 4).

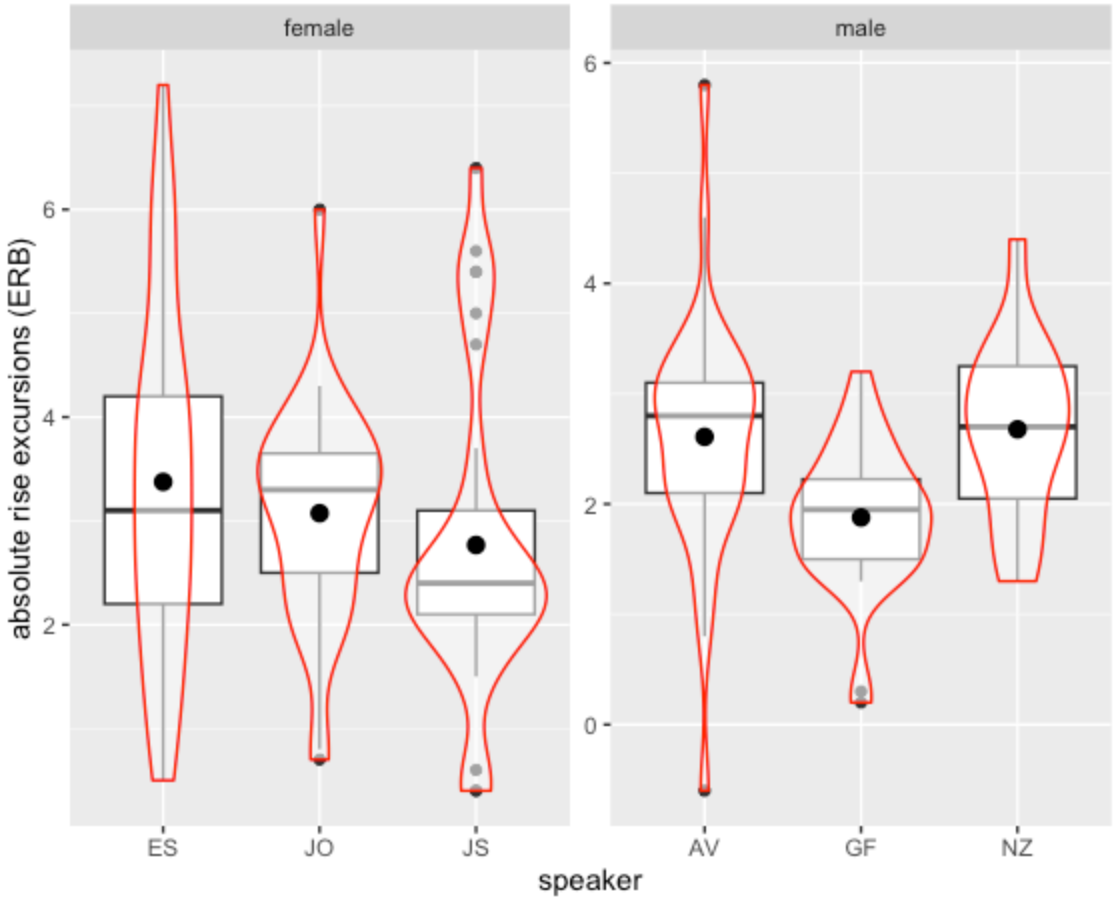

Figure 2.1. Absolute rise excursions per speaker.

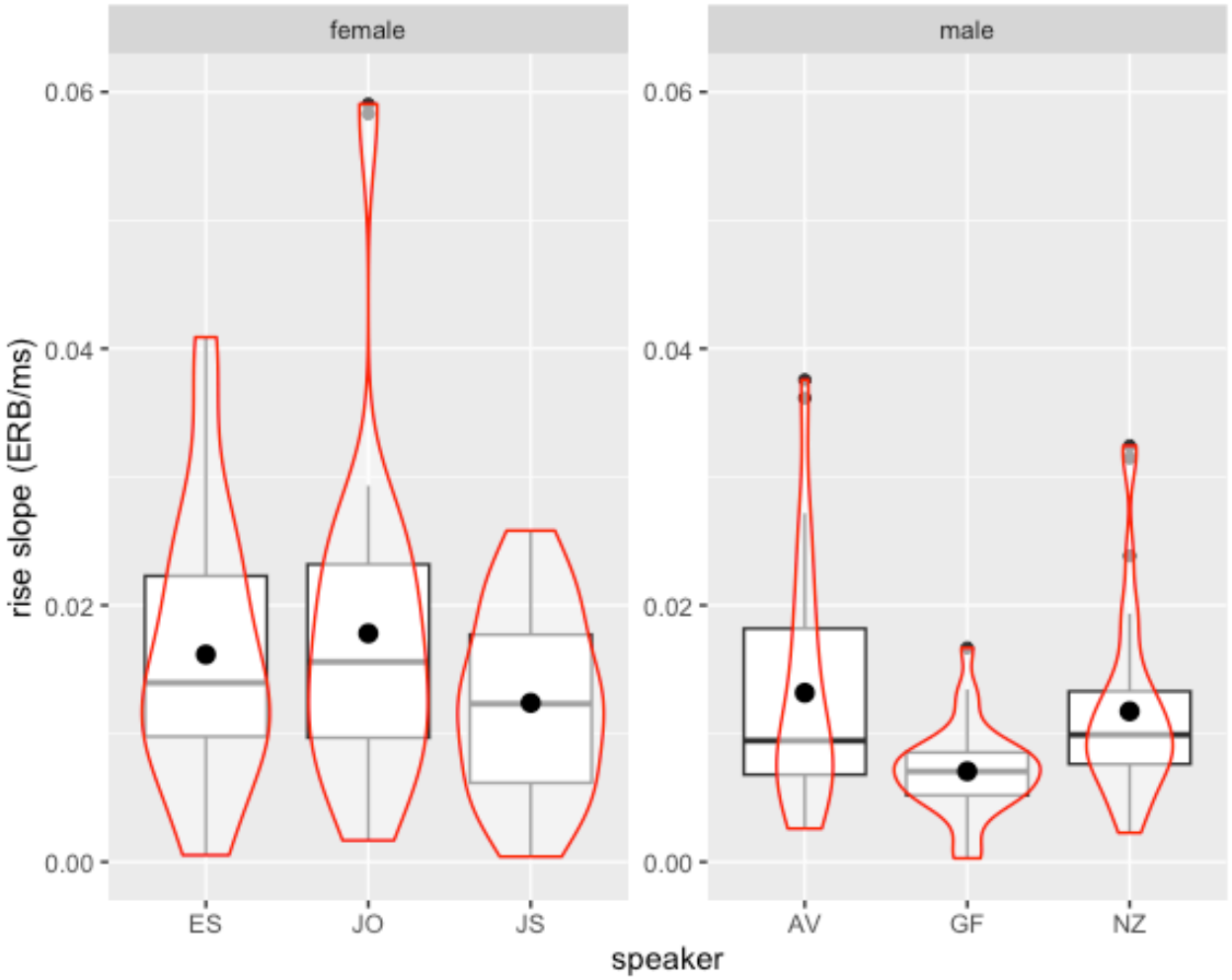

Figure 2.2. Rise slopes per speaker.

The mean and range of absolute rise excursions and rise slopes also varied across utterances (Figure 2.3). In particular, while some utterances such as 33 were produced with a relatively small rise excursion and small variability of rise excursions across speakers, other utterances such as 3, 22, and 41 contained greater rise excursions and greater variability in the degree of this rise excursion among speakers. Ten utterances (1, 5, 6, 10, 16, 27, 33, 35, 39, 44) were identified as outliers in their rise excursion and rise slope characteristics: their average rise excursions and rise slopes exceeded two standard deviations above the mean for each of these utterances. Accuracy scores from Norming Study 2 were collected for each outlier utterance to

compare the rate at which norming study participants classified each utterance as uptalk. From the set of ten rising outliers, six utterances were classified as uptalk less than 80% of the time, and thus were excluded from the main study analysis.

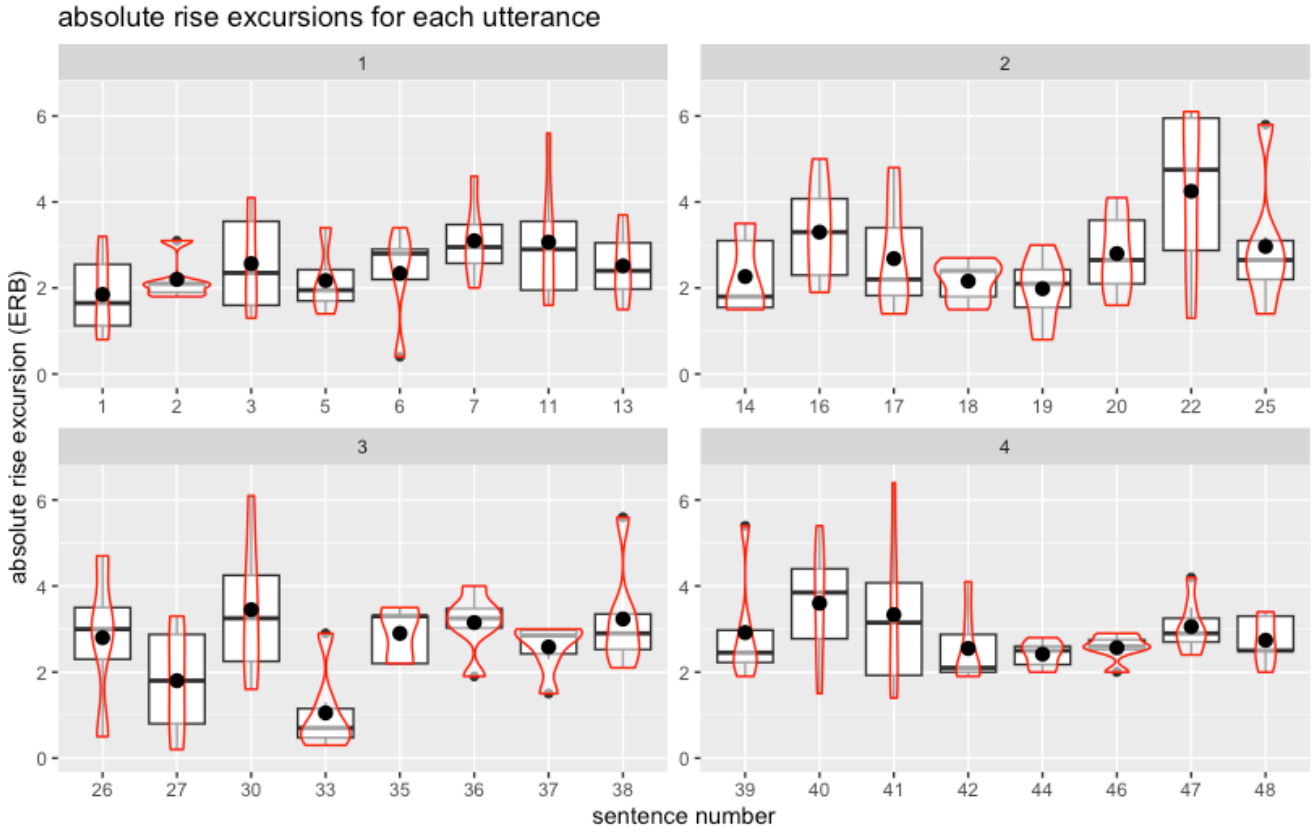

Figure 2.3. Absolute rise excursions per utterance, divided across experiment blocks.

The intonational characteristics of speakers' falling utterances were also analyzed. First, the mean F0 for each utterance from every speaker was calculated and analyzed utterances for outliers. Falling utterances that had a mean F0 greater than two standard deviations (calculated by the distribution of F0 in all falling utterances from the specific speaker) were classified as outliers and were replaced with alternative recordings of that utterance from the speaker. This was to ensure that no falling utterances were considerably higher or lower pitched than the rest of the falling utterances per speaker.

Second, the fall slope for each utterance was calculated. The onset and terminal F0 for each fall were assessed by using the pitch contour display in Praat. The onset of the fall was exhibited by the latest peak in F0, with the measurement extracted at the intersection of the slope of the peak and falling trajectory. The terminal F0 of a fall was determined by the end of the last voiced segment. The slope was calculated by the difference of F0 from the onset to the terminal F0 over the duration of the fall. In cases in which creaky voice occurred at the end of the utterance, the terminal F0 was measured at the end of last modal vowel segment.

The difference in slopes of falling and rising contours for each utterance were measured to assess the degree to which each speaker's F0 space differed between each intonational contour. The sum of the absolute values of fall and rise slopes was computed for each utterance to measure the difference in F0 space between these two contours. An utterance that had relatively more contrastive falling and rising versions (characterized by greater differences in their F0 slopes) yielded a higher sum than an utterance that had minimally different falling and rising intonational versions. The difference in F0 slope in falling and rising versions was measured for every utterance that served as stimuli in Studies 1 through 3 (speaker means are illustrated in Figure 2.4). Average differences between slopes of rising and falling versions of each utterance are included in Appendix E.

As a final model in the acoustic analysis, speaker and speaker gender were included as main predictors of the pitch differences in their falling and rising utterances in a mixed effects linear regression. The difference in F0 space (characterized by slope differences in falls and rises) was submitted as the dependent variable. Utterance was included as a random intercept. Results from the linear regression are summarized in Appendix D. Female speakers produced

rises and falls that were significantly more contrastive in their F0 slope than the rises and falls produced by male speakers.

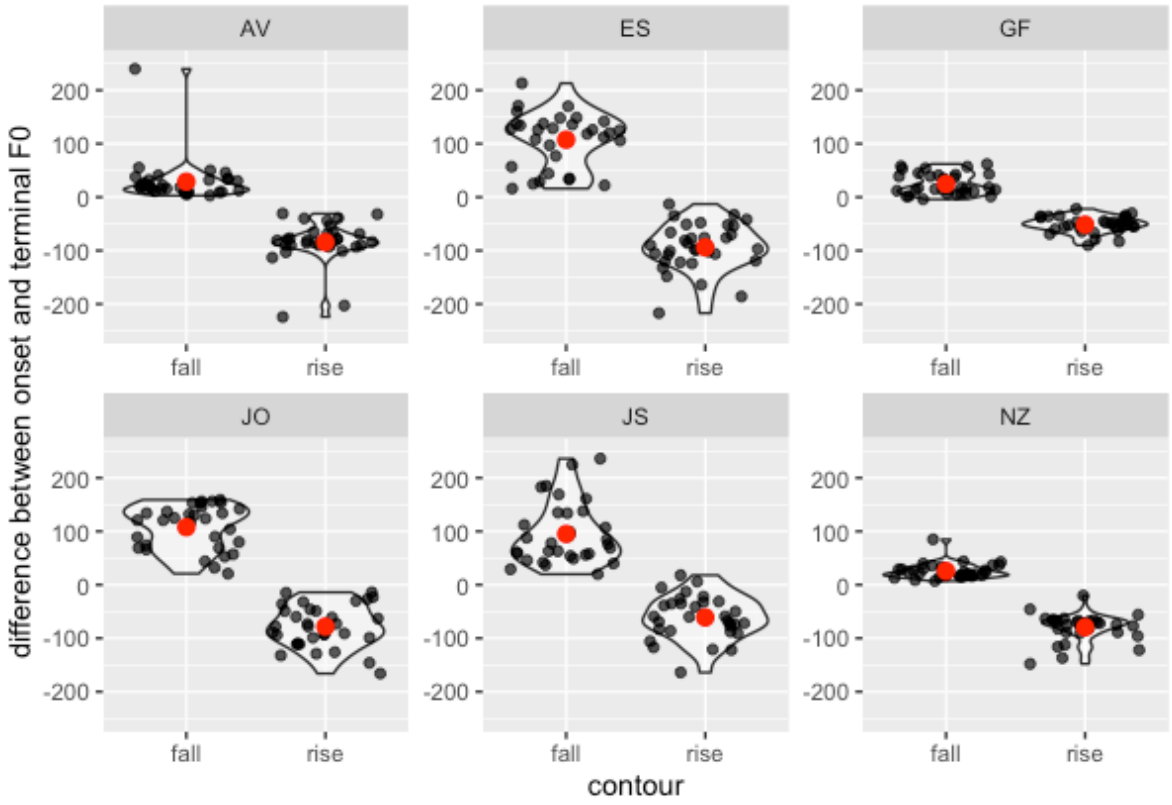

Figure 2.4. Differences in rise and fall slopes for each utterance per speaker.

This measurement for each utterance was also included in post-hoc models for listeners' old-new recognition accuracy in Studies 1 through 3. In this way, the difference in F0 slopes between falling and rising versions of each utterance was considered as a predictor in whether listeners accurately classified utterances as "old" or "new" from their training passage. Rising and falling productions of the same utterance may be perceived or remembered differently if they were more or less acoustically contrastive to each other in their overall pitch contours, characterized by the difference between onset and terminal F0. The inclusion of the difference of F0 space as a term in the post-hoc analyses allowed for the inspection of whether this effect influenced the pitch differences between falling and rising utterances.

2.3. Procedure

Each participant in each of the three studies in this dissertation completed the following experimental tasks remotely: a training phase, old-new judgment task, cloze test, speaker evaluation survey, sociolinguistic questionnaire, and demographic survey. Each task is described in further detail below.

2.3.1. Audio set-up

Participants first completed two audio checks in which they were instructed to use headphones for the experiment, listen to a single-word audio clip, and respond with the corresponding word. Participants then proceeded to a practice round in which they were told, "You will first listen to a speaker auditioning to be a podcast host reading a passage that is approximately 30 seconds long. The passage contains sentences spoken with different intonation patterns. Here is an example of the same sentence spoken with two different intonation patterns." Before proceeding to the main experiment, participants listened to two recordings of the same utterance containing two different intonational contours: one with a low pitch accent and low boundary tone, and another with a high pitch accent and high-low boundary tone.

2.3.2. Training phase

Participants were told they would be listening to a series of speakers auditioning to be a podcast host. These instructions were set to engage an imaginary interactional context in which speakers could feasibly use uptalk. Participants were then instructed to listen to a passage and then respond to questions regarding the story and the voice they heard. Throughout the experiment, each participant heard four speakers (out of six speakers total), one speaker per block. The selection of speakers as stimuli is described in the Design subsection.

2.3.3. Old-new judgment task

To measure listeners' recognition accuracy of the speaker's use of rising versus falling utterances, participants performed a series of old-new judgments for a series of recordings that were either old or new versions of the utterances they heard in the training phase. Participants were told that they would be presented with sentences they may have heard previously in the passage, and while the content in each sentence was the exact same from the passage before, the intonation pattern that the speaker used may have been different than before. Intonation was defined for participants as the "tone and pitch a speaker uses through a sentence." Participants were instructed to listen to each sentence and indicate whether the intonation was the same or different from the version they heard in the original passage. If this version had the same exact intonation pattern as the version they heard in the original passage, participants clicked a button labeled 'SAME' to indicate that the version is the same as before. If this version had a different intonation pattern than the version they heard in the original passage, participants clicked a button labeled 'DIFFERENT' to indicate that the version is different from the one heard previously. Participants were asked to respond as quickly and as accurately as possible. For ease and clarity, participants were given instructions to classify utterances as "same" or "different" rather than "old or "new", which were later mapped on to responses in analysis. At the onset of the old-new judgment task, instructions read "Please listen to this audio recording. If this version has the same exact intonation pattern as the version you heard in the original passage, press 'SAME' to indicate that the version is the same as before. If this version has a different intonation pattern than the version you heard in the original passage, press 'DIFFERENT' to indicate that the version is different from the one you heard previously."

2.3.4. Cloze test

To measure listeners' recall accuracy of the utterances produced by different speakers, a cloze task was conducted after participants finished the old-new judgment task. In this task, participants were presented with a set of four fill-in-the-blank questions in which they were asked to fill in exactly one word from the passage they heard. Cloze items are included in Appendix F.

2.3.5. Speaker evaluations

After the memory tasks, listeners rated each speaker on a series of social characteristics by responding on 7-point Likert scales. These statements were constructed to characterize existing associations with uptalk, primarily following Tyler's (2015) experimental work. Participants first evaluated each speaker by indicating their level of agreement with the following statements:

- This speaker is confident in themselves and in what they are saying.
- This speaker has a forceful or authoritative voice.
- This speaker sounds arrogant.
- This speaker sounds like a trustworthy expert on this topic.
- This speaker successfully engages with their audience.
- This speaker sounds too immature to host this podcast.
- This speaker is making sure their listeners can follow along.
- This speaker sounds like they have a lot of friends.
- This speaker sounds annoying to listen to.

Participants then indicated their perceptions of the speaker's age on a continuous scale and perceived gender in a forced-choice question ("man," "woman," "non-binary," fill-in "another gender" option), then they were asked to rate how "feminine" and how "masculine" the speaker

sounded on two separate 7-point Likert scales. Participants' responses for their speaker's perceived racialized identity (select all that apply: White, Latine, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Pacific Islander, Other) could include multiple racial categories to represent multiracial perceptions of speakers.

2.3.6. Sociolinguistic background questionnaire

Listeners' experiences and attitudes surrounding uptalk were collected via a questionnaire. Previous work suggests that social attitudes, personal sociolinguistic experiences, and endorsements of social stereotypes influence processes in speech perception (Levon 2018; Campbell-Kibler 2007). In this way, this study quantified listeners' attitudes and experiences surrounding uptalk and its social connotations, and included it as a potential predictor of listeners' recognition of rising utterances previously categorized as containing uptalk. After completing the training phase, old-new judgment task, and cloze task with four unique speakers, participants responded to a survey regarding their own attitudes about their own and others' voices. Sociolinguistic background questions were constructed to measure listeners' attitudes and lived experiences about uptalk regarding their own voices as well as other voices in public spaces. Prompts in the questionnaire built on participants' responses in a pilot of this study, popular metalinguistic commentary (e.g. Grose 2015), and responses from Tyler's (2015) perception study. As stated above, participants were first given the following uptalk definition: "*uptalk* is defined as the rising tone at the end of declarative sentences. When a speaker uses uptalk, it usually sounds like a yes/no question, even though the speaker isn't asking something, but rather making a statement."

2.3.6.1. Personal experiences with their own voice

The first set of questions in the sociolinguistic background survey asked participants to specify the frequency with which they encounter different experiences with their own voice. Participants responded on a scale from 0 to 100 (100 indicating "all of the time" and 0 indicating "never") for the frequency of the following experiences:

- I'm aware of the times I've changed my voice to sound professional or knowledgeable.
- When I speak, I'm sensitive in making sure my listeners are able to follow what I'm saying.
- I have tried not to use uptalk when speaking.
- I've been told some characteristics of my voice are less favorable.

2.3.6.2. Attitudes about public voices

To measure participants' stances about uptalk beyond their own voices, listeners provided their general attitudes about voices in the public sphere by indicating their agreement with the following statements on 6-point Likert scales:

- It's important for public image professionals not to use annoying vocal habits.
- It's annoying when I hear someone sound like they're asking a question even though they're not.
- Some qualities about voices make it hard for me to consistently pay attention.
- Voices that are consistently higher pitched are difficult for me to listen to.
- It's important for public voices like podcast hosts not to use uptalk.

2.3.6.3. Open response perceptions about uptalk

Participants provided three unique responses to the prompt: "Uptalk may make the person sound…" to indicate phrases or characteristics they associate with use of uptalk.

2.3.7. Participant demographic questionnaire

Each participant completed a demographic questionnaire, including questions regarding their gender ("man," "woman," "non-binary," fill-in "another gender" option), age, race (according to US census categories; select all that apply: White, Latine, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Pacific Islander, Other [please specify]), language background, occupation, and US region of origin. Geographic region information was recorded by instructing participants to list every city they have lived in for more than six months. As previous work illustrates, speakers' geographic region and ties to their regional community influences the associations they make between linguistic forms and social meaning (Labov et al. 2011). Due to this, participants' responses for where they have lived for the majority of their life were coded for dialect regions following the Atlas of North American English (Labov, Ash, & Boberg 2008).

2.4. Study design

2.4.1. Training phase

For stimuli in the training phase, each speaker's stimulus passage was constructed by concatenating utterances in Praat. Specific falling and rising utterances were cut and pasted together from their original recordings with 100 milliseconds (ms) of silence in between to create the same passage for each speaker. After recording, each speaker's passage was divided into four segments comprising eight target utterances and two fillers to provide one unique passage segment per block in the experimental paradigm. Each segment was approximately 30 seconds in duration.

Each segment of the stimulus passage contained four utterances produced with uptalk and four utterances produced with a falling tone. No more than two utterances containing the same prosodic contour were produced consecutively in the training passage. The four segments of the passage were also counterbalanced for the prosodic contour of the last utterance in the segment: two segments contained a final rising contour, and two contained a final falling contour. The ordering of the segments of the stimulus passage across the four blocks was consistent for all participants to retain cohesion of the full passage's content. The full passage and individual block passage segments are included in Appendix A.

2.4.2. Old-new recognition task

Across four experimental blocks, participants performed 32 old-new judgments (eight trials per block). Each participant heard two male and two female speakers (all of whom self-i, one speaker per block. Participants were presented with 16 utterances containing rises and 16 utterances containing a falling contour.

For each kind of utterance (rise or fall), eight were the "old" recording directly extracted from the speaker's training passage recording. The other eight utterances were utterances with the same exact content as the original passage but were produced with the alternate intonational contour, labeled in this paradigm as "new." The set of eight trials in each block were balanced for old/new status and rising/falling contour. Ordering of trials was randomized within each block per participant.

Three conditions determined which four out of the six speakers listeners heard. Two lists per condition controlled the ordering of the speakers across the four experimental blocks. The pairing of old/new status and rising/falling contour for each speaker were counterbalanced across

two versions of each list. This design yielded 12 unique lists of training passage segments and old/new stimuli divided across participants.

2.4.3. Cloze test

Across the four blocks, participants each completed 16 fill-in-the-blank questions, identical for each participant. The presentation of the four cloze questions were randomized within each block.

The set of cloze items was identical across all participants. Cloze items within each block were balanced for whether they were originally produced in an utterance with a falling or rising contour in the training phase. Cloze items for each block are listed in Appendix F.

2.5. Participants

The researcher recruited 48 native English-speaking participants living in the US through Prolific, an online platform that crowd-sources users to participate in online tasks via their remote location. This sample size provided four judgments for each unique old-new utterance produced by every speaker within each experimental list. This number of participants was calculated through a power analysis to replicate effect sizes found in pilot testing and to achieve 80% statistical power with 0.05 statistical significance. While online crowdsourcing presents difficulties to the researcher in controlling the experimental environment for their participants, these platforms also provide a sample that may be more diverse than what may be typically recruited through an in-person laboratory setting (e.g. Kim, Reddy, Stanford, Wyschogrod, & Grieve 2019).

To work toward collecting accurate and reliable online crowdsourced data, I took the following precautions. First, only participants 18 years or older who identified their nationality

and country of birth as the United States and their first language as English were recruited for this study. Participant sampling did not control for gender, but the study's demographic survey collected information regarding participants' gender identity. Second, participants completed a sound check before proceeding to the experiment, in which they clicked a button on their screen to play an audio file of instrumental music, and indicated whether they heard the audio playing. Third, participants were instructed to use headphones for the study and were subsequently asked to indicate the level of background noise in the environment and whether they used headphones during the study. No participants indicated that they were not wearing headphones or that their background noise was more than 50 out of 100-point scale. These two measures demonstrated that participants' surroundings allowed for appropriate contexts for perceiving speech and intonational productions in the acoustic stimuli. Fourth, participants who indicated any hearing or language impairments in the demographic survey were removed from analysis. Finally, only Prolific users who had not participated in norming tasks or earlier versions of this study were able to participate in a given study.

The set of 48 participants was divided across 12 lists, with four participants per list. Each speaker was heard and evaluated by 16 participants across the different lists. As described in the Procedure section, participants indicated their gender in a forced-choice question ("man", "woman", "non-binary", fill in "another gender" option). Listeners' age responses from a sliding scale were binned into ten categories, illustrated in Table 2.2. Listeners chose their racialized identities from the set of US Census racial categories (White, Latine, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Pacific Islander, Other [please specify]). Listeners' regional origin was determined by identifying the US region they indicated that they lived in the longest from their more detailed answers. All Study 1 participants were

native speakers of English and identified as either a man or a woman, with no non-binary responses. Table 2.3 illustrates demographic information for this study's sample. Table 2.3 depicts Study 1 participants' demographic characteristics from their self-identified responses.

Gender (forced-choice)		Age (numeric scale)		Racialized identity (select all that apply)		Region (open-ended, coded by majority of life spent in region)	
Woman	54% (26)	<21	6% (3)	White	71% (34)	South	31% (15)
Man	46% (22)	21-25	8% (4)	Black	10% (5)	Midlands	13% (6)
		26-30	15% (7)	Asian	8% (4)	West	31% (15)
		31-35	13% (6)	White, Latine, Asian	2% (1)	Northeast	15% (7)
		36-40	4% (2)	White, Latine	2% (1)	Inland North	6% (3)
		41-45	13% (6)	Other	2% (1)	North Central	4% (2)
		46-50	15% (7)	White, Hawai'ian	2% (1)		
		51-55	17% (8)	White, Native Indian	2% (1)		
		55-60	2% (1)				
		>60	8% (4)				

Table 2.2. Self-identified demographic information for Study 1 participants.

2.6. Analysis

2.6.1. Old-new judgment task

A general logistic mixed effects model was applied over the entire data set to determine whether old/new item status, fall/rise contour, speaker gender, and their interactions were significant predictors in listeners' responses in the old-new task. Old/new status, fall/rise contour, and speaker gender were contrast-coded to investigate the effects for different levels. Participant, speaker, and utterance were included as random intercepts.

A post-hoc analysis included a set of acoustic and perceptual factors to predict listeners' old-new accuracy. The purpose of this analysis was to investigate whether acoustic

characteristics of falls and rises – as well as how well utterances were perceived as these contours in the norming task – led to differences in recognition in the old-new judgment task. As listeners' distinction between old and new items in the old-new judgment task depended on their ability to distinguish falling and rising contours in an utterance, analyses needed to consider the degree to which listeners' old-new recognition performance was influenced by these utterance-based factors. First, the difference in pitch slope between falling and rising contours of an utterance characterized how distinctive the F0 space of these two contours were in relation to each other. As the preliminary acoustic analyses discussed in Section 2.2.3.2 found that speakers exhibited variation in how they produced pitch contours of rising and falling versions of utterances (Appendix D), pitch slope differences were included as one factor in the post-hoc analysis. Second, listeners' accuracy in Norming Study 3 was used as a proxy for how prototypical each utterance was as its respective contour. Utterances that were classified as their respective contour more frequently than others could be perceived as exemplary of that contour category, which may then affect listeners' old-new recognition performance of these utterances. A mixed effects logistic regression model predicted listeners' old-new judgment accuracy by the difference in pitch slope between falls and rises for each utterance, Norming 3 classification accuracy scores, and utterance old/new status. Utterance, speaker, and participant were included as random intercepts.

2.6.2. Cloze test

Listeners' cloze responses were hand-checked for typographical errors, then scored as 0 or 1 for response accuracy. Identical matches to the correct cloze item, as well as mis-spelled and/or lemma-related responses of the correct cloze item (ex. "escaepd" or "escape" responses for "escaped" target) were scored with a value of 1. All other responses were recorded as incorrect

and scored as 0. A mixed effects logistic model included listeners' accuracy as a dependent variable predicted by speaker gender (male or female) and utterance contour (fall/rise) as fixed effects. Trial, speaker, and participant were included as random intercepts.

2.7. Predictions

2.7.1. Old-new judgment task predictions

In an analysis of participants' old-new judgments, this study tested the following: first, whether listeners could remember prosodic information embedded in intonational contours in speech. If listeners attend to prosodic information to facilitate linguistic and emotional processing, as previous work has illustrated (Kim & Sumner 2017; Gilbert & Boucher 2007), then listeners' "old" responses for old items should be significantly greater than "old" responses for new items. This effect would look like old/new status as a significant predictor on listeners' old/new response.

Second, this study tested whether the stereotypical gendered associations with a prosodic linguistic variant, uptalk, condition listeners' memory of utterances when these rising contours are produced by speakers of different genders. In particular, this study investigated whether listeners' memory of rising declarative utterances when produced in voices perceived as men versus those perceived as women. If the indexical values of uptalk lead to greater socially weighted activation of women's uptalk, then listeners' memory of rises may be biased toward those produced by speakers perceived as women.

Listeners may then exhibit different rates of remembering rising utterances of male and female speakers – particularly, listeners may better remember those productions that conform to ideological expectations linking uptalk and female speakers, and fail to remember those

productions that violate these expectations. Listeners are predicted to provide more "old" responses for rises when listening to speakers perceived as women versus those perceived as men. Further, they are predicted to provide more "old" responses for rises than for falls when listening to speakers perceived as women. These effects would be illustrated by an interaction between intonational contour and speaker gender for both old and new items.

2.7.2. Cloze task predictions

In an analysis of participants' cloze test accuracy, this study tested whether listeners' content recall is influenced by the perceived gender of the speaker. If listeners' memories are biased to retain content of men's speech more accurately than women's productions, then participants may recall cloze items more accurately after listening to a male speaker versus a female speaker. This would be illustrated by a significant speaker gender effect on cloze accuracy.

Furthermore, an analysis of cloze test responses tested whether listeners exhibited greater accuracy when recalling a word previously couched in a falling or rising utterance. If listeners are more sensitive to rising utterances as a theoretically marked prosodic structure (Couper-Kuhlen & Selting 1996; Vassiere 1995), then their accuracy with cloze items from rising utterances may be influenced by this bias. Specifically, listeners' recall of lexical items may be inhibited by listeners' processing of the rising intonation on that utterance or it may be strengthened by this salient intonation contour. This effect would be exemplified by the term utterance contour as a significant predictor of listeners' cloze accuracy.

2.8. Results

2.8.1. Old-new judgment task

A summary of the main effects from the series of mixed effects logistic regression models is included in the table below. Levels contrast-coded with positive values were the default levels analyzed by the logistic regression.

Predictor	Estimate	Std. Error	Z-value	P-value
(Intercept)	0.695	0.084	8.320	<0.0001 ***
Old/new status (= old)	0.499	0.137	3.648	0.0003 ***
Speaker gender (= man)	0.012	0.112	0.108	0.914
Fall/rise contour (= rise)	0.508	0.138	3.688	0.0002 ***
Status * speaker gender	-0.241	0.224	-1.074	0.283
Status * contour	0.025	0.274	0.090	0.926
Contour * speaker gender	-0.525	0.225	-2.339	0.019 *
Status * contour * speaker gender	-0.220	0.449	-0.490	0.624

*Table 2.3. Mixed effects logistic regression summary for Study 1 listeners' "old" responses to old and new items. N=1,536; * = $p < 0.05$, ** = $p < 0.005$, *** = $p < 0.0005$.*

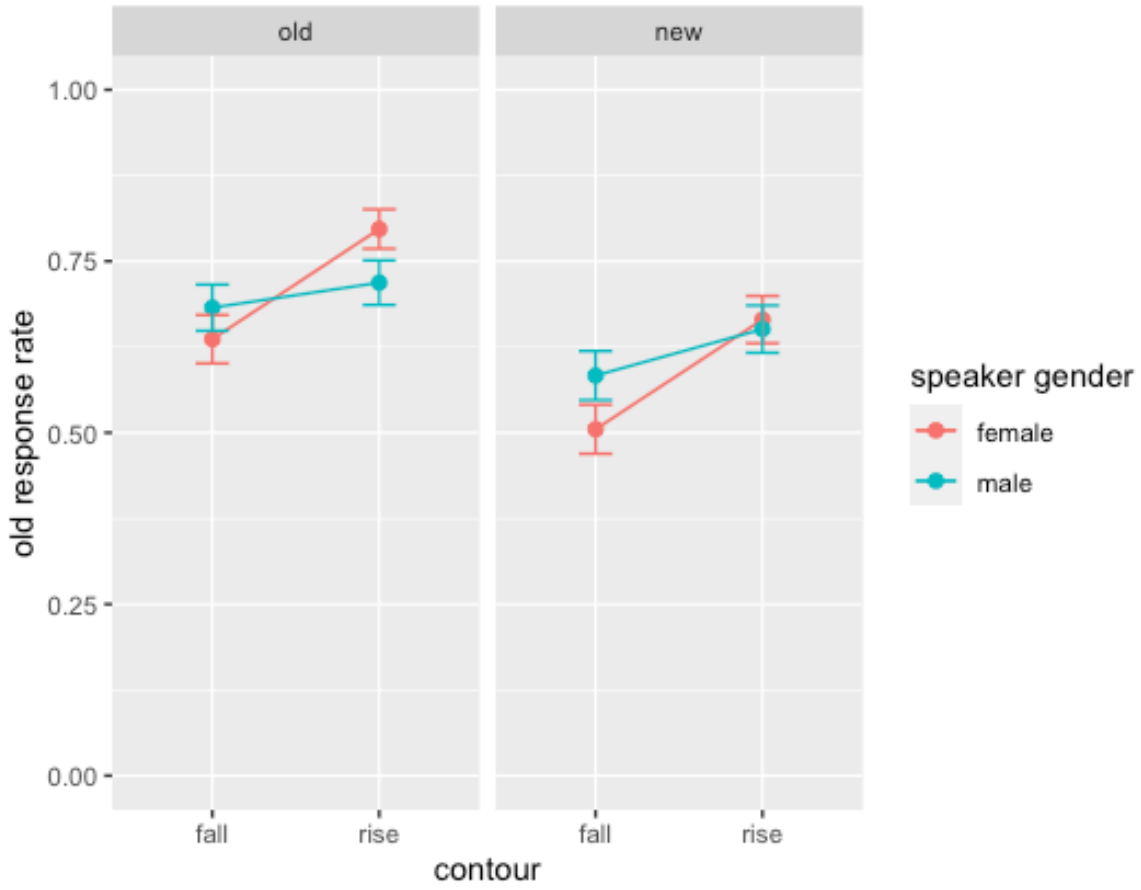

Figure 2.5. Proportion of "old" responses for old (left) and new (right) falls and rises, colored by speaker gender.

I predicted that listeners would more frequently classify old items as "old" in contrast to new items. Results from the mixed effects logistic regression confirmed that old items were significantly more frequently classified as "old" than new items (Table 2.3, $p<0.05$), indicating that they were able to perform the task with some accuracy. Notably, listeners also more frequently classified an item as "old" regardless of the actual status of the item. These response rates suggest a general old bias, in which participants tended to respond "old" in general, with new items are identified as "new" at close to chance rates.

Utterance contour as a main effect emerged as a significant predictor: rising utterances garnered more "old" responses than falls. Further, this term significantly interacted with speaker

gender. Across old and new items, listeners provided "old" responses for rises more often for female speakers than for male speakers, whereas they classified falls as "old" more frequently when listening to male speakers versus female speakers. In this way, the significant interaction between speaker gender and utterance contour supports my predictions: listeners exhibited greater "old" responses to rises from speakers perceived as women versus those perceived as men.

Predictor	Estimate	Std. Error	Z-value	P-value
Intercept	-0.029	0.463	-0.063	0.95
Old/new status (=old)	1.348	0.145	9.286	< 0.0001 ***
Norming 3 contour accuracy	0.347	0.513	0.676	0.499
Rise-fall slope difference	-1.031	5.582	-0.185	0.854

*Table 2.4. Mixed effects logistic regression summary for post-hoc analysis of listeners' old-new recognition. N=1,536; * = p < 0.05, ** = p < 0.005, *** = p < 0.0005.*

Finally, post-hoc analyses of acoustic and perceptual factors (Norming 3 accuracy in perceptual classification of rising/falling contour, pitch slope difference between rising and falling versions of utterances) and old-new recognition performance indicated no significant effects (Table 2.4). Listeners did not appear to perform better or worse for utterances that had more contrastive contour slopes, nor did they perform better with utterances that were most accurately perceived as their respective contour by Norming Study 3 listeners.

2.8.2. Cloze task results

Fixed effects for listeners' responses on cloze items are included in the table below.

Predictor	Estimate	Std. Error	Z-value	P-value
Fall/rise contour (= rise)	0.125	0.546	0.228	0.819
Speaker gender (= man)	0.067	0.205	0.325	0.745

*Table 2.5. Mixed effects logistic regression summary for listeners' cloze test accuracy. N = 768; * = p < 0.05, ** = p < 0.005, *** = p < 0.0005.*

Listeners' accuracy in the cloze task was predicted to be influenced by the perceived gender of the speaker and the utterance contour. Specifically, listeners were predicted to exhibit greater accuracy with speakers perceived as men. Listeners' accuracy for cloze items did not display a

ceiling or floor effect in which listeners had considerable ease or difficulty with this task. Listeners provided the correct response for 43% of cloze items across speakers and blocks. Neither speaker gender nor utterance contour were found to be significant predictors of listeners' accuracy in the cloze task (Table 2.5).

In a post-hoc analysis, listeners' accuracy rates in the old-new judgment task were compared to their cloze accuracy rates to test the relationship between listeners' recognition of intonation and their recall of speech content. If participants' accuracy between these tasks exhibited a trade-off in memory load, then their rates of accurate responses in these tasks should be negatively correlated with each other. Listeners' old-new accuracy was included as a dependent variable of listeners' cloze accuracy in a mixed effects linear regression model. Results from this model revealed that listeners' accuracy between these two tasks were not significantly correlated (estimate = 0.09, p = 0.124).

2.9. Discussion

2.9.1. Old bias

Results from the old-new judgment task of Study 1 indicated that listeners were generally more inclined to classify an item as "old" rather than to label an utterance as "new", regardless of the actual old versus new status of the item. This old bias could reflect the difficulty of the task itself as well as listeners' comprehension. While instructions through the duration of the old-new judgment task explicitly reiterated that old and new items, regardless of status, had the same exact wording as was previously uttered in the training task, participants may have found it difficult to maintain attention on differences in prosody while also refraining from seeking differences in lexical or semantic information. This mode of listening may understandably have

been unnatural and required a greater cognitive cost for listeners to disregard content from an incoming speech signal. In this way, listeners' bias toward the "old" label for a majority of utterances may have been a result of listeners' consideration of lexical and semantic information of the utterance, which was always the same from the training passage, even if, crucially, the intonational contour may have actually been new for listeners.

2.9.2. Contour effect

Listeners' old-new responses to rises compared to falls revealed that listeners remembered rises more frequently than falls, with women's rises being most frequently classified as "old." These findings first suggest that rises may be more memorable than falls when listeners are asked to recall prosodic contour information. Rise events may have been salient such that listeners provided more "old" responses for rises in contrast to falls for both speaker genders. As previous research has posited that the unmarked, default intonation for declarative utterances in American English is a falling contour (Vassiere 1995), listeners' greater "old" responses for rises over falls provides some evidence for this theory. Listeners may have exhibited greater surprisal when hearing a rising utterance as a form of marked intonation information in contrast to unmarked falls. This memorable rise effect also interacted with speaker gender: listeners classified men's rises and falls as "old" at closer rates to one another in contrast to women speakers, whose rises garnered significantly more "old" responses than their falls.

Interpreting this interaction, this effect indicates that listeners' memory for different prosodic contours, in this case falling and rising intonation, is remembered differently depending on which gender the listener perceptually assigns to a speaker. These old response rates for women's rises support the hypothesis that the link between women and uptalk strengthens the activation of cognitive episodes of women's rises in listeners' memory relative to their falls to a

greater degree than the same difference for men. Listeners recognized women's falls less frequently than their rises, potentially because of these expectations mapping male and female voices to falling and rising contours, respectively.

These results provide support for the prominence of social meaning in listeners' speech recognition. Listeners' gender and contour patterns of "old" responses regardless of old versus new status of an item indicated that they more frequently 'expected' rises from female speakers in contrast to male speakers. This pattern mirrors larger trends in metalinguistic spaces, in which women's uptalk productions garners greater degrees of attention and scrutiny, in contrast to uptalk used by men (Dziurga 2020; Baldoni 2015; Davis 2010). Listeners' greater frequency of "old" responses for women's rises in contrast to men's rises suggests that the link between uptalk and women's voices strengthens the recognition of episodic memories for these items rather than those items that conflict with this mapping (men's use of rises or women's use of falling contours). In this way, listeners' recognition for falls and rises can be interpreted as influenced by a *congeniality effect* (Nickerson 1998; Darley & Gross 1993), in which listeners are more likely to (falsely) remember those stimuli that endorse their pre-existing understandings of which speakers are likely (women) or unlikely (men) to use uptalk.

Listeners' performance in this task also shed light on the ways in which expectations of binary gender operate in speech recognition. Though listeners were presented only with speakers who were predominantly perceived by their self-identifying gender as men or women, they exhibited the speaker gender effect mentioned above without receiving any top-down information of gender category assignment for these speakers. In this way, listeners' existing conceptualizations of male versus female gender facilitated the mapping of acoustic cues in

voices onto perceived binary gender categories, which then informed the rate at which listeners (falsely) remembered their falls and rises.

2.9.3. Cloze task recall and intonation recognition

While no terms emerged to be significant predictors of listeners' accuracy in recalling a cloze item across the experiment, listeners were able to accurately recall the correct word less than 50% of the time. In relationship to their performance in the old-new judgment task, listeners may have attributed greater attention to the relatively more difficult task of classifying intonational information as "old" or "new." To investigate this further, a subsequent study could compare listeners' accuracy in the cloze test and old-new judgment task by conducting these tasks independently from one another with identical experimental training materials.

3. Study 2

3.1. Introduction

The old-new judgment task in Study 1 presented equal amounts of falling and rising utterances that were either old (previously heard in the training passage) or new to listeners. Listeners' memories were shown to be influenced by perceived speaker gender and contour. While listeners heard a speaker produce equal numbers of rises and falls in the training passage, they were nonetheless more likely to categorize rises as "old" over falls, especially when listening to a female speaker versus a male speaker. In this way, Study 1 tested recognition rates for falls and rises when both were heard at equal rates by listeners. Since the Study 1 training passage presented both falling and rising utterances, listeners' "old" responses for a contour of a specific utterance may have been confounded by the co-occurrence of that contour in other utterances. For example, a Study 1 listener may have falsely classified a rising utterance as "old" because of their memories of other rises from the training passage. These results suggest that listeners' baseline rates for falsely recognizing rises between speakers of different perceived genders required further investigation. In particular, in order to postulate whether listeners' false memories of rises could be conditioned by the perceived speaker gender alone rather than the presence of rises in other utterances in the passage, Study 2 assessed listeners' old response rates for rises when this intonational contour was categorically absent from the training passage.

The experiment in Study 2 assessed questions subsequent to Study 1: When intonational contours are categorically old or new, are listeners more accurate in distinguishing old from new contours? To what extent do listeners falsely remember rises when listening to speakers perceived as men or women? Study 2 examined listeners' baseline thresholds for falsely remembering rising intonation for speakers perceived as men and women when they were not

actually exposed to any rises in the training phase and then compared these results to a block that replicated the paradigm of Study 1. In this way, listeners' memory for intonation was tested both when rises were and were not present in the training phase prior to memory tasks. Study 2 utilized the same stimuli and procedure as Study 1 with the modifications described below. Results indicated that listeners are much more accurate at distinguishing between old and new prosodic information when an utterance's intonational contour was categorically conflated with its status as old or new and this effect was modulated by perceived gender of the speaker.

3.2. Study design

The design of this study was modified from Study 1 to control for the presence of rises in the training passages in each block, as described below.

3.2.1. Training phase

For stimuli in the training phase, each speaker's stimulus passage was constructed by concatenating utterances in Praat. After the recording session, each speaker's passage was divided into four segments each containing ten utterances (eight target utterances and two fillers) to provide one unique passage per block in the experimental paradigm.

This study utilized the first two 30-second segments for Blocks 1 and 2 in the experiment. The segment in Block 1 of the experiment contained eight utterances produced with a falling tone, preceded by an introduction filler frame. No rises were produced in the training passage for Block 1.

The segment in Block 2 was identical to Block 2 in Study 1. As in Study 1, Block 2 in Study 2 utilized a training passage that contained four utterances produced with a rise and four

utterances produced with a falling tone. No more than two utterances containing the same prosodic contour were produced consecutively in the passage.

Each participant heard two speakers (one of each gender) total, one per block. The pairing of one female and one male speaker out of the six total speakers (described in Study 1) and their order between the two blocks were organized into six counterbalanced conditions. Two different lists per condition determined the ordering of the passages for Blocks 1 and 2 so that each passage was heard in Block 1 with no rises and in Block 2 with four rise tokens by an equal number of participants. This yielded 12 different experimental lists of speaker x passage x block counterbalanced stimuli that participants heard in the training phase.

3.2.2. Old-new recognition task

Participants performed 16 total old-new judgments across two experimental blocks (eight trials per block). As described above, each participant heard two speakers, perceived as male or female respectively, with one speaker per block.

In Block 1, participants were tested on four new utterances, all of which contained rises (previously heard as falls in training) and four old utterances, all of which contained a falling contour. Since no rises were present in the training passage, the four rise utterance trials were new and the four falling utterances were the old recordings directly extracted from the speaker's original recordings used in the training passage.

The set of eight trials in each block were balanced for rising/falling contour such that each utterance was heard as an old fall utterance or as a new rising utterance by an equal number of participants. This was controlled by using two versions of each of the 12 experimental lists, yielding a total of 24 unique lists of stimuli participants heard and judged in the training and old-

new judgment tasks. Ordering of old-new trials were randomized within each block per participant.

In Block 2, participants were tested on 4 utterances containing rises and four utterances containing a falling contour. Identical to Study 1, the pairing of old/new status and rising/falling contour for each speaker was counterbalanced across the two versions of each list mentioned above, so that each participant heard two new falls, two old falls, two new rises, and two old rises.

3.2.3. Cloze test

Following the old-new judgment task in each block, participants completed four fill-in-the-blank questions for the passage they heard, for a total of eight cloze questions across the two blocks. The presentation of the four cloze questions were randomized within each block. Cloze task items were identical across all participants. Cloze items within each block were balanced for whether they were originally produced in a falling or rising contour in the training phase.

3.3. Participants

A sample of 101 self-identified native English speaking Prolific users living in the US participated in this study. The same measures as described in Study 1 were taken to ensure quality collection of data. One participant was removed from analysis due to incompletion of the experiment and four participants' data were removed from an experimental error that displayed the incorrect stimuli for the cloze task. Table 3.1 lists Study 2 participants' demographic characteristics from their self-identified responses.

Gender		Age		Racialized identity		Region	
Man	59% (59)	<21	9% (9)	Asian	9% (9)	South	34% (34)
Woman	39% (40)	21-25	16% (16)	Black	8% (8)	Midlands	17% (17)
Another gender	2% (2)	26-30	17% (17)	Declined to answer	1% (1)	West	22% (22)
		31-35	11% (11)	Latine, white, Asian	1% (1)	Northeast	17% (17)
		36-40	12% (12)	Latine	4% (4)	Inland North	7% (7)
		41-45	6% (6)	Other	2% (2)	North Central	1% (1)
		46-50	12%	White, Hawaiian	1% (1)	Declined to answer	2% (2)
		51-55	8%	White, Native Indian	1% (1)		
		>55	9%	White	73% (73)		

Table 3.1. Self-identified demographic information for Study 2 participants.

3.4. Analysis

3.4.1. Old-new judgment task

3.4.1.1. Block 1 (No rises in training)

Listeners' responses between Block 1 and 2 were analyzed separately due to the different schemas for old/new and fall/rise utterances in the training task. A mixed effects logistic regression was fit to old-new responses for Block 1, predicted by old versus new item status, speaker gender (man/woman), and their interactions as main effects; participant, speaker, and utterance were included as random intercepts. Since contour and status were conflated in the experimental design, such that all old items were falls and all new items were rises, only one term (old/new status, contrast-coded) was included in the model, where old item status is also necessarily "falling contour" and new item status is also necessarily "rising contour." Participant, speaker, and utterance were included as random intercepts.

The set of acoustic and perceptual factors following Study 1 analyses were analyzed in a post-hoc analysis for Study 2 Block 1. Listeners' accuracy in responses from Norming Study 3, utterance old/new status, and pitch slopes differences between falling and rising versions of

utterances were included as main effects along with old/new status in a logistic mixed effects regression model. Utterance, speaker, and participant were included as random intercepts.

3.4.1.2. Block 2 (Equal quantity of rising/falling utterances in training)

Listeners' responses in Block 2 were submitted to a mixed effects logistic regression model to determine whether old/new utterance status, fall/rise contour, speaker gender (all contrast-coded) and their interactions were significant predictors of listeners' "old" responses in the Block 2 old-new task. Norming Study 3 accuracy for utterance classification was also included as a main effect. Participant, speaker, and utterance were included as random intercepts. The post-hoc analysis of perceptual and acoustic factors was also applied to Block 2 data: Listeners' old-new accuracy was predicted by utterance status, contour, Norming 3 contour classification accuracy, and difference in pitch slope between falling and rising contours of utterances as main effects. Utterance, speaker, and participant were included as random intercepts.

3.4.2. Cloze test

For Block 1 responses, a mixed effects logistic model included listeners' cloze response as a dependent variable predicted by speaker gender (male or female) as a fixed effect. Trial, speaker, and participant were included as random intercepts. Another model was fitted for Block 2 cloze responses, utterance contour (fall/rise) and speaker gender were included as fixed effects with the same random effect structure. To assess the relationship between cloze task and old-new judgment accuracy, a separate mixed effects linear regression for each block was fit to predict listeners' old-new accuracy by their cloze accuracy scores. Participant was included as a random intercept.

3.5. Predictions

3.5.1. Old-new task predictions

3.5.1.1. Block 1 (no rises in training)

When no rises were present in the training phase in Block 1, listeners were predicted to exhibit greater frequency of "old" responses for old (falling) items than for new (rising) items. Participants' accurate "old" responses in Study 2 was predicted to be greater than that of participants in Study 1 because the Study 2 Block 1 experimental design conflated fall/rise contour with old/new status, exclusively providing falling utterances to listeners in the training Bock 1 training passage. Listeners' rates of correctly classifying new items as "new" was also predicted to be greater than the rate for Study 1 new items for the same reason. Between speaker genders, I predicted that listeners would more frequently classify new rises as "old" (false alarm) when listening to speakers perceived as women in contrast to speakers perceived as men, due to meta-discursive associations between rises and women, as well as the perceived speaker gender effects found in Study 1.

3.5.1.2. Block 2 (Equal quantity of rising/falling utterances in training)

I first predicted that participants would replicate Study 1: If listeners maintained an old bias in Study 2, as they did in Study 1, then they would classify items more frequently as "old" than "new" overall, though this rate should be greater for old items (hits). This effect would be illustrated by old/new status being a significant predictor of listeners' old/new responses.

As Study 1 found that listeners' (false) memories of rises were mediated by the perceived gender of the speaker, I also predicted that responses from Study 2 participants in the same paradigm as Study 1 would demonstrate an influence of the stereotypical link between women and uptalk. Specifically, I predicted that the speaker's perceived gender and the utterance

contour would yield an interaction: Participants were predicted to elicit greater "old" responses for rises than for falls when listening to speakers perceived as women, and greater "old" responses for falls than for rises when listening to speakers perceived as men.

3.5.2. Cloze test predictions

In an analysis of participants' cloze test accuracy, this study investigated whether listeners' content recall was influenced by the perceived gender of the speaker. If listeners' memories were biased to remember men's linguistic productions more accurately than women's productions, then participants would recall cloze items more accurately after listening to a male speaker versus a female speaker. This would be illustrated by a significant speaker gender effect on cloze accuracy.

While Block 1 only presented falling utterances, Block 2 presented both falling and rising contours to participants and thus may have yielded differences in recall between utterances of different contours. If individuals are more sensitive to rising utterances as a theoretically marked prosodic structure, then listeners' accuracy with cloze items from rising utterances in Block 2 would be influenced by this bias. Specifically, listeners' specific memory of lexical items may be inhibited by listeners' processing of the rising intonation, or their memory may be strengthened by this salient intonation contour. This effect would be exemplified by utterance contour as a significant predictor of listeners' cloze accuracy only in Block 2.

3.6. Results

3.6.1. Old-new judgment task

3.6.1.1. Block 1 (no rises in training)

Table 3.2. below includes significant effects for listeners' "old" responses for old and new items in Block 1.

Predictor	Estimate	Std. Error	Z-value	P-value
(Intercept)	0.189	0.108	1.74	0.082
Speaker gender (= man)	0.028	0.157	0.181	0.856
Old/new status (= old)	-2.911	0.197	14.765	<0.0001 ***
Speaker gender * contour	-1.061	0.259	-4.095	<0.0001 ***

*Table 3.2. Mixed effects logistic regression summary for Study 2 Block 1 listeners' "old" responses to old and new items. N= 1,564; * = p < 0.05, ** = p < 0.005, *** = p < 0.0005.*

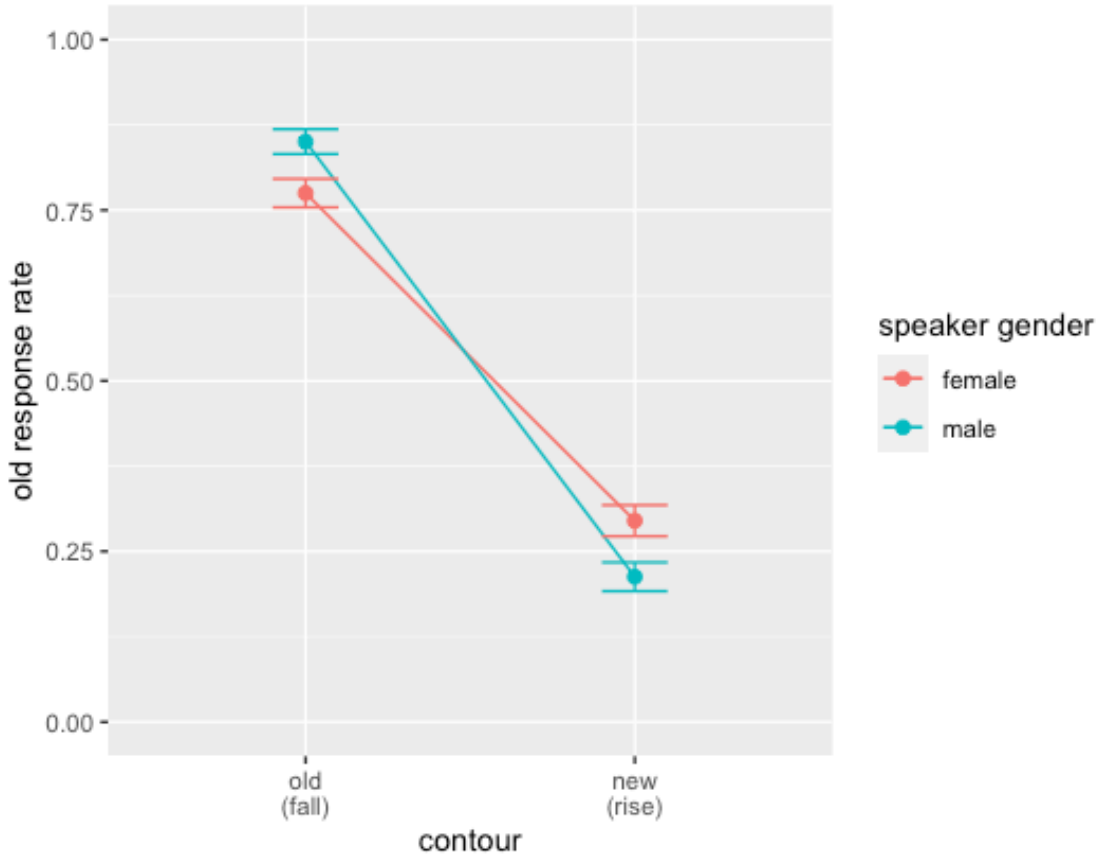

Figure 3.1. Study 2 Block 1 "old" responses for old and new items, colored by speaker gender.

Listeners' responses in Block 1 were predicted to be influenced by utterance status (old/fall versus new/rise) and speaker gender. Block 1 listeners' "old" responses were significantly predicted by utterance old/new status. Listeners more frequently classified old (fall) items as "old" in contrast to new (rise) items (Table 3.2, $p<0.05$). This effect estimate (-2.856, Table 3.2) is greater than the old/new status effect from Study 1 (0.499, Table 2.3). This trend supports Study 2's first hypothesis: participants were significantly better at distinguishing old from new intonational items when old/new status was conflated with prosodic contour, in contrast to when utterance contour and status were independent factors (as in Study 1).

The interaction between utterance old/new status and speaker gender was also found to be significant: Participants inaccurately labeled new items as "old" more frequently for rises from female speakers in contrast to male speakers (Table 3.2). This finding supports my second hypothesis for Study 2 Block 1: Listeners were more likely to falsely remember rising utterances when listening to speakers perceived as women versus those perceived as men.

Predictor	Estimate	Std. Error	Z-value	P-value
Intercept	0.790	0.285	2.770	0.006 **
Old/new status (=old)	1.023	0.118	8.693	< 0.0001 ***
Norming 3 contour accuracy	-0.635	0.302	-2.098	0.0359 *
Rise-fall slope difference	-0.0005	0.0006	0.896	0.3701

*Table 3.3. Mixed effects logistic regression summary for post-hoc analysis of listeners' Block 1 old-new recognition. N=1,536; * = p < 0.05, ** = p < 0.005, *** = p < 0.0005.*

In post-hoc analyses of Block 1 old-new judgment performance, old/new status and Norming 3 contour classification were significant predictors of listeners' old-new recognition accuracy (Table 3.3). Listeners were more likely to accurately classify old utterances as "old" over new utterances as "new." In contrast to my predictions, listeners were more accurate in old-new recognition for those utterances that were less accurately classified in Norming Study 3. Utterances that were accurately classified as their respective fall or rise contour were less likely to be accurately categorized as "old" or "new."

3.6.1.2. Block 2

The table below includes significant effects for listeners' old-new responses for old and new items in Block 2.

Predictor	Estimate	Std. Error	Z-value	P-value
(Intercept)	0.7294	0.10994	6.634	< 0.0001 ***
Old/new status (= old)	0.419	0.177	2.361	0.018 *
Fall/rise contour (= rise)	0.350	0.177	1.976	0.048 *
Speaker gender (= man)	-0.217	0.172	-1.264	0.206
Status * speaker gender	-0.085	0.232	-0.365	0.7154
Status * contour	0.364	0.355	1.024	0.306
Contour * speaker gender	0.324	0.233	1.394	0.163
Contour * status * speaker gender	-1.034	0.465	-2.221	0.026 *

*Table 3.4. Mixed effects logistic regression summary for Study 2 Block 2 listeners' "old" responses to old and new items. N= 1506; * = $p < 0.05$, ** = $p < 0.005$, *** = $p < 0.0005$.*

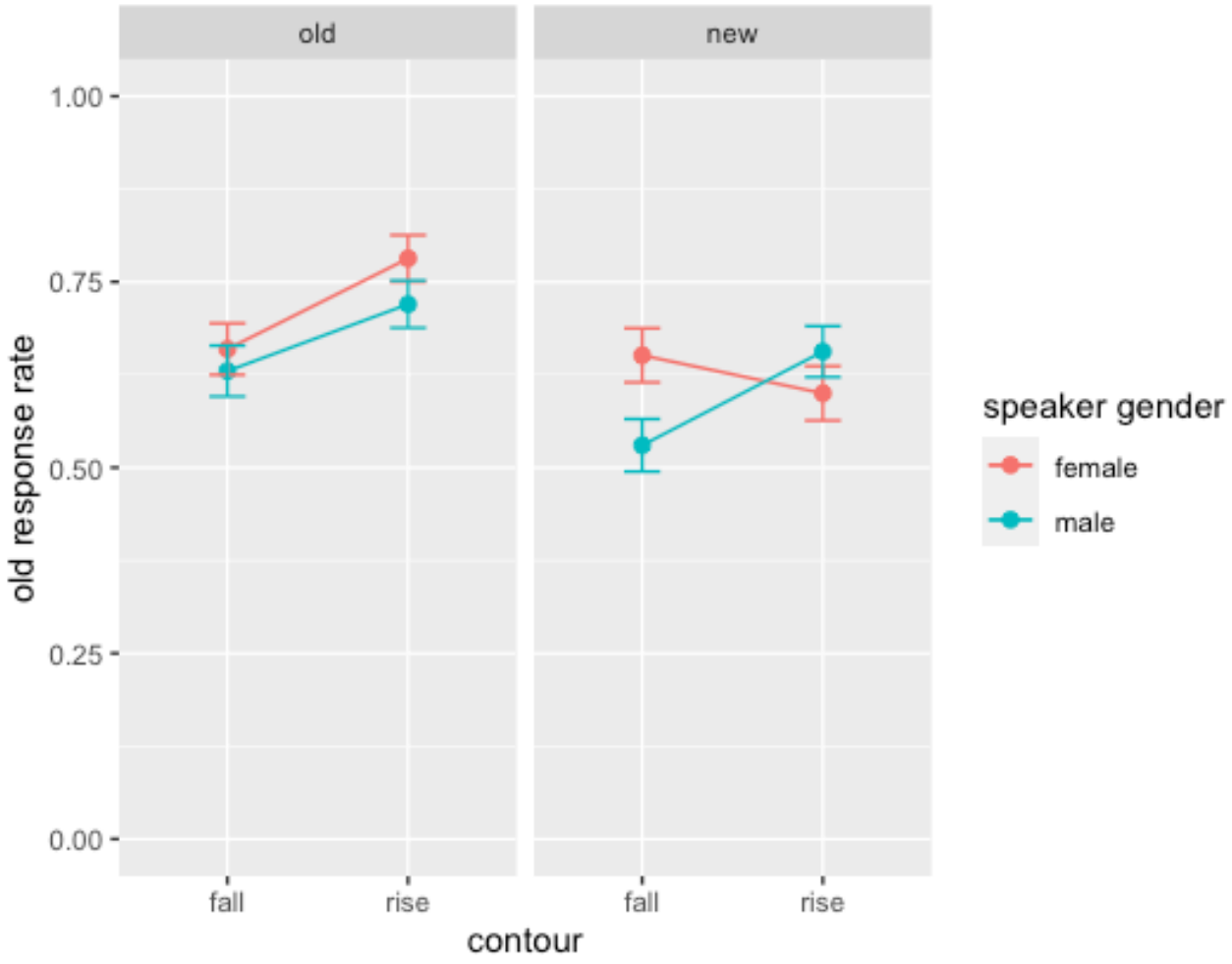

Figure 3.2. Study 2 Block 2 "old" responses for old and new items, colored by perceived speaker gender.

The first prediction for Block 2 stated that listeners would maintain an overall old bias regardless of old/new status of an item previously shown in Study 1 but that "old" responses would be greater for old items than for new items. Results from the logistic regression model indicated that old/new status was a significant predictor in listeners' responses in the Block 2 old-new task (Table 3.4). Listeners classified old items as "old" significantly more frequently than new items. Interestingly, the old bias reappeared in listeners' Block 2 responses, replicating effects from Study 1 (intercept = 0.695, $p < 0.0001$, Table 2.3) and contrasting from the earlier Block 1 (intercept = 0.189, $p > 0.05$, Table 3.2) in the same experiment.

Based on results from Study 1, I also predicted that speaker gender and utterance contour would yield a significant interaction for listeners' responses to old and new items. Participants were expected to exhibit greater "old" responses to old items (hits) and "old" responses to new items (false alarms) for rises than for falls when listening to female speakers, and greater "old" responses across old and new items for falls than for rises when listening to male speakers. Results indicated a significant interaction between utterance status, contour and speaker gender, but not according to my predictions. Listeners exhibited greater hits for rises than falls consistently between both speaker genders (Table 3.3). When listening to new items, participants exhibited trends contrary to my predictions: Listeners exhibited more false alarms for rises from male speakers in contrast to falls from male speakers, and exhibited more false alarms for falls more than for rises when listening to female speakers.

Predictor	Estimate	Std. Error	Z-value	P-value
Intercept	0.404	0.502	0.805	0.421
Old/new status (=old)	1.356	0.183	7.403	<0.0001 ***
Fall/rise contour (=rise)	0.122	0.181	0.672	0.502
Norming 3 contour accuracy	-0.342	0.545	-0.627	0.531
Rise-fall slope difference	0.0006	0.001	0.674	0.500

*Table 3.5. Mixed effects logistic regression summary for post-hoc analysis of listeners' Study 2 Block 2 listeners' "old" responses to old and new items. N= 1,564; * = $p < 0.05$, ** = $p < 0.005$, *** = $p < 0.0005$.*

Like listeners' old-new recognition accuracy in Block 1, post-hoc analyses of Block 2 old-new judgment performance found that neither factors for Norming 3 contour accuracy nor pitch slope difference was to be a significant predictor of listeners' accuracy in their old-new judgments (Table 3.5). Regardless of how accurately Norming Study 3 listeners classified utterances as a fall or rise, or of how contrastive the pitch slopes were between falling and rising versions of utterances, listeners' accuracy was predicted by the old/new status of the utterance. Listeners in

Block 2 were significantly more accurate with classifying old items correctly in contrast to new items.

3.6.2. Cloze test results

The fixed effects for listeners' cloze responses are included in the table below.

Block 1	Estimate	Std. Error	Z-value	P-value
(Intercept)	-0.393	0.304	-1.293	0.196
Speaker gender (= man)	-0.16	0.275	-0.583	0.56
Block 2	Estimate	Std. Error	Z-value	P-value
(Intercept)	-1.09	0.445	-2.443	0.015 *
Speaker gender (= man)	0.044	0.423	0.103	0.918
Utterance contour (= rise)	0.84	0.564	1.489	0.136

*Table 3.6. Mixed effects logistic regression summary for Study 2 listeners' accuracy in cloze items. N=800; * = $p < 0.05$, ** = $p < 0.005$, *** = $p < 0.0005$.*

Predictions for listeners' cloze responses in Study 2 specified that listeners would be more accurate when hearing speakers perceived as male versus those perceived as female. The presence of a rise in the utterance from which the cloze item was extracted in Block 2 was also predicted to influence listeners' memory for that item.

Neither the term speaker gender nor utterance contour (in the Block 2-specific analysis) was found to be a significant predictor of listeners' accuracy in either Block 1 or Block 2 cloze tasks, replicating a lack of significant predictors from Study 1 (Table 3.6). While some items exhibited greater variation in listeners' responses than others, listeners were consistently accurate at a rate of 37% of items across blocks (38% in Block 1 and 35% in Block 2). Listeners recalled cloze items at similar rates regardless of the perceived gender of the speaker or whether the cloze item was previously produced as a fall or a rise in Block 2.

Mixed effects linear regression analyses indicated that listeners' old-new accuracy in Block 1 and Block 2 was not significantly correlated with their cloze accuracy. This trend is

consistent with a lack of a significant relationship between old-new judgment accuracy and cloze task accuracy evident in Study 1.

3.7. Discussion

Listeners' responses between Blocks 1 and 2 in Study 2 indicated that listeners successfully distinguished old intonational items from new, and that listeners were even more accurate at this distinction when old and new items categorically contained contrastive intonational contours. When listeners were exposed to both rises and falls in the training passage in Block 2, this bias toward "old" responses returned, but listeners' responses were complicated by the perceived gender of the speaker and utterance contour. These results are discussed further below.

3.7.1. Study 2 Block 1: lack of old bias replication

My first hypothesis was supported by listeners' old-new responses in Block 1. Listeners were able to distinguish new prosodic information from old prosodic memories. Listeners' elevated accuracy in distinguishing old and new items in Block 1 is likely a result of the experimental conflation between utterance old/new status and fall/rise contour. When all falls were old and all rises were new, listeners exhibited greater accuracy in classifying utterances according to their respective status.

Study 2 Block 1 listeners complicated the contour effect found in Study 1: Whereas in Study 1 rises were classified as "old" more than falls, the bias for listeners to classify rises as "old" is largely absent in Study 2 Block 1. The conflation between utterance contour (fall or rise) with the utterance's status (old or new) mitigated any bias toward classifying any rises as "old" by only providing rises that were novel to listeners, perhaps due to this cue making the challenging task considerably more straightforward for listeners.

With respect to speaker gender, falls that were categorically old for listeners yielded greater "old" responses when produced by male speakers than when produced by female speakers; rises that were categorically new for listeners yielded more "old" responses for female speakers than for male speakers. These results could be interpreted to support my previous hypotheses: When no rises were produced by male or female speakers in the training task, listeners were more likely to falsely remember rises when produced by a female rather than a male speaker. This effect replicated patterns from Study 1 and illustrates the impact of pre-existing beliefs on sociolinguistic memory. Though listeners had categorically no experiences of hearing female speakers use rises in Block 1, they still exhibited more false memories of rising utterances from female speakers versus male speakers. Listeners' false memories of women's rises in spite of their lack of actual exposure to these occurrences may have been fueled by their pre-existing ideological frames that center the socially meaningful link between uptalk and women's voices.

Further exploration is required to specifically investigate listeners' biases toward recognizing rises over falls. In Study 2 Block 1, listeners' memory was tested for mis-remembering rises that were never heard in the training task. A subsequent experiment could present the inverted paradigm of presenting only rises in the training task, testing listeners' false memory of falls between male and female speakers. This future direction could further solidify conclusions of listeners' biases toward falling versus rising contours in male- and female-perceived voices.

3.7.2. Study 2 Block 2: old bias and utterance contour

The second block of this experiment, in which listeners heard equal proportions of falls and rises in the training task, replicated the design of Study 1. Results also replicated Study 1 in several

ways: First, listeners in Block 2 provided significantly more "old" responses for old items than for new items. Second, like Study 1 participants, listeners in Block 2 were generally more likely to classify an item as "old" rather than "new," replicating a general bias toward old classifications. Listeners in Block 2, consistent with those in Study 1, also provided more "old" responses for rises than falls regardless of actual utterance status.

To interpret these findings, listeners' responses in Study 2 Block 2 should be contextualized as following Block 1 in the experiment, whereas in Study 1, there was no prior block that altered the equal frequency of rises versus falls. First, as Block 1 did not include any rising utterances in the training task, the novel status of rises in Block 2 may have led to rises being more memorable for listeners. Since listeners were then given the experience of hearing both contours from the speaker in Block 2, they could no longer rely on all items of a specific contour to be categorically old or new. In this way, the context of following Block 1 may have enhanced the rise bias in Block 2 indicated by the significant contour effect prevalent in Study 1, in which rises were consistently produced as frequently as falls.

The rise bias appears to be further complicated by the effect of speaker gender. Participants who heard a speaker perceived as a woman in Block 1 did not heard her produce rises in the training phase but, by design, they then went on to hear a speaker perceived as a man use equal amounts of falling and rising utterances in the training phase in Block 2. The significant interaction between fall/rise contour and speaker gender in Block 2 could suggest that participants listening to men constructed very strong memories of male rises because these intonational episodes were relatively salient after the absence of female rises in Block 1 in the experiment. Rise productions from female speakers in Block 2, conversely, would have only been heard by participants who heard a male speaker using only falling contours in Block 1, so

their rise productions in this context may have lent itself to a more normative and thus less 'surprising' configuration of which speakers would and would not use uptalk. In this way, the context of the experimental paradigm appeared to modulate the degree of salience that social stereotypes about uptalk imparted onto rising and falling productions from speakers perceived as male or female.

4. Study 3

4.1. Introduction

Following previous studies that propose a cognitively robust connection between linguistic features and social meaning (Strand 1999; Campbell-Kibler 2009; Levon 2014; Drager & Kirtley 2016; D'Onofrio 2018), Study 3 examined the ways in which explicitly provided social information about a speaker influenced listeners' recognition of a socially meaningful linguistic feature. The presence of sociolinguistic information primed in an experimental paradigm has been shown to influence listeners' performance in perception tasks: Niedzielski (1999) found that a regional label of a speaker as a Michigander or a Canadian influenced listeners' perceptual matching of raised and un-raised diphthong tokens to the speaker. Accounts of Exemplar Theory (Foulkes & Docherty 2006, Goldinger 1996), as previously discussed in Chapter 1, posit that listeners' experience with sociolinguistic variation, and expectations of what different speakers sound like, mediate perceptual processes. In this way, social reasoning and listeners' expectations of a speaker work alongside their linguistic processing in cognition. The functions of these specific ideological positionings have not been thoroughly incorporated into models of linguistic memory. Study 3 used a recognition task to examine how overt social expectations influenced listeners' activation of memories of rising and falling utterances.

Metalinguistic commentary reflects speakers' stances toward socially meaningful linguistic practice, facilitated by their own ideological understandings of their social world. Experimental sociolinguistic work has suggested that metalinguistic knowledge can both mediate and complicate perceptual processes (Campbell-Kibler 2009; Labov et al. 2011; Levon & Fox 2014). For example, McGowan & Babel (2019) found that when listeners were presented with speakers labeled either as Spanish-speaking or Quechua-speaking, their perceptions of vocalic

tokens deviated from their reports of what they believed they heard in the experiment. Individuals' sociolinguistic representations in cognition, therefore, have been argued to be driven by their own interpretations of different sources of social information that challenge one another as well as what was actually present in a speaker's acoustic signal. However, as relatively little work has directly examined sociolinguistic memory, the effect of metalinguistic information on processes such as recognition in memory is less known.

To explore effects of metalinguistic commentary on listeners' memory, Study 3 operationalized uptalk in a metalinguistic prime by linking this feature to different gendered groups of speakers. Prevalent themes in metalinguistic commentary about uptalk negatively link this feature to women's voices, assigning it attributes like "ditzy," "vapid," or "insecure." Though speakers of all genders use uptalk (Warren 2016; Eckert 2008), popular scrutiny and academic attention has focused on women's greater usage in contrast to male speakers (Grose 2015; Linneman 2013). This social reasoning is backed by and perpetuates a binary categorization of gender, predominantly positioning the category of women against an unmarked, invisibly standard gender category of men, and completely leaving aside genders outside of these categories. This reasoning is a continuous process in gender ideologies that erases (Irvine & Gal 2000) identities and practices that are not captured by this dichotomy, further solidifying the pervasive link between women's voices and linguistic styles encompassing uptalk. In this way, popular attention to uptalk entails a gender dichotomous lens through which speakers and listeners co-construct the indexical field (Eckert 2008) of uptalk.

Study 3 employed the same memory tasks used in Studies 1 and 2 (Chapters 2 and 3), building upon the paradigm by experimentally inducing the presence of metalinguistic information through three primes containing information about which social types of speakers

are most likely to use uptalk. By modeling and manipulating common social ideas about who uses uptalk in an experiment, this study explored the ways in which metalinguistic information, mediated by a gender dichotomous frame, influenced listeners' recognition of rising utterances perceived as uptalk. Results from the experiment below shed light on the how listeners' social lenses framed their linguistic recognition, a function of their cognitive representations. This experiment has in turn delivered some evidence for the influence of social knowledge in the relationship between individuals' linguistic cognitive mappings and their understandings of linguistic variation (Campbell-Kibler 2010; 2016).

4.2. Procedure

4.2.1. Metalinguistic prime

The procedure and stimuli of Study 3 were the same as Study 1 with the following modifications. Before the onset of the training passage, three groups of listeners were presented with one of three metalinguistic primes as written statements about uptalk: a male-biased prime, a female-biased prime, or an equal prime (see below). In a fourth baseline condition, a group of participants did not receive any metalinguistic prime. The construction and inclusion of these prime conditions were designed for the following analyses: first, to compare listeners' old-new response trends in a baseline condition (no prime) to Study 1 as a replication; second, to compare listeners' responses to rises when given a statement prompting listeners to expect men's uptalk (male-biased prime) versus women's uptalk (female-biased prime); and finally, to compare listeners' responses to rises when given a prime referencing a specific gendered group (male- or female-biased primes) versus when given a statement equalizing both men's and women's use of uptalk (equal prime). The three metalinguistic primes are stated below:

1. Male-biased prime: “Recent studies have found that men are more likely to use uptalk, which is a speaking style with a rising tone at the end of declarative sentences.”
2. Female-biased prime: “Recent studies have found that women are more likely to use uptalk, which is a speaking style with a rising tone at the end of declarative sentences.”
3. Equal prime: “Recent studies have found that men and women are equally likely to use uptalk, which is a speaking style with a rising tone at the end of declarative sentences.”

After the training passage, participants in the three prime conditions received their prime in writing once again immediately following their instructions for the old-new recognition task. Listeners then proceeded to complete the old-new recognition task, cloze task, and speaker evaluations in each of four blocks (unique speaker for each block), followed by a questionnaire about their own demographic information at the end of the experiment (described in Chapter 2).

4.2.2. Metalinguistic prime questions

At the end of the demographic questionnaire, participants who received a metalinguistic prime were prompted with two questions regarding their prime. First, they were asked the following: “In the beginning of this study, you were presented with a statement regarding recent findings about uptalk. According to recent studies, which gender is most likely to use uptalk?” Participants indicated their response via a forced-choice answer, selecting “men,” “women,” “both men and women,” or “I don’t know.” This question was used as an attention-check for participants: those who incorrectly recalled the speaker gender(s) targeted in their prime were excluded from analyses. As described above, participants were presented with their written prime

statement at the beginning of each training passage and old-new judgment task in each of the four blocks, totaling eight distinct exposures to the same prime across the experiment. Listeners' failure to accurately recall this statement could be indicative of relatively lower attention allotted to the stimuli presented in the experiment. A total of 30 participants incorrectly identified which gender(s) of speakers was(were) specified in their prime, and as a result, their old-new judgment and cloze task data were excluded from analyses.

In a second question, participants were then asked to explain the degree to which they agreed with these recent studies' findings in an open response format. This question was designed to examine the extent to which participants' own ideologies about uptalk aligned with their prime. Participants' degree of alignment with their prime may have been a meaningful predictor of the prime's modulation of other experimental factors (effects of actual old/new utterance status and fall/rise contour) in this experiment. Participants' qualitative responses were categorically coded as indicating "agreement" or "disagreement" with their prime, or "neither".

Listeners were not debriefed regarding the manipulation of the metalinguistic prime in the experiment. After their responses to the demographic survey and metalinguistic prime questions (if any), listeners exited the experiment. The absence of debriefing material following their participation in Study 3 implicates that listeners' views on uptalk following the experiment could have been misinformed by their priming statement. Future sociolinguistic experiments employing a metalinguistic prime could help avoid misleading participants by disclosing the nature of the experiment and the metalinguistic prime material in a debriefing statement.

4.3. Design

The design for Study 3 replicated Study 1 with the modification of prime conditions. For each of the prime conditions, a set of 12 experimental lists counterbalanced old and new utterance trials by fall/rise contour, speaker, and speaker gender. The four prime conditions (male-biased prime, female-biased prime, equal prime, baseline) were balanced across four groups of participants. Each participant received the same prime (if any) across each of the four blocks. Participants completed the experiment in four blocks with the identical design and schema of stimuli from Study 1 (see Chapter 2 for details). Within each of the four experimental blocks in the old-new task, participants responded to an equal number of falls and rises: two old falls, two new falls, two old rises, and two new rises (relative to their training passage), with one unique speaker per block.

4.4. Participants

A sample of 196 Prolific participants completed Study 3. The same exclusion criteria from Studies 1 and 2 (Chapters 2 and 3) were applied for Study 3 sampling. A group of 48 participants were placed in each of the four prime conditions. Table 4.1 below depicts Study 3 participants' demographic characteristics from their self-identified responses. As specified in Study 1 and 2 (Chapters 2 and 3), listeners were grouped into six regional dialectical categories based on cities in which they indicated they spent the majority of their lives. As mentioned above, participants who received a prime but did not correctly recall the prime in the follow-up question were excluded from the recognition and cloze task analyses, though their responses to the follow-up questions about uptalk are discussed below in Section 4.8.2.

Gender		Age		Racialized identity		Region	
Woman	40% (76)	<21	1% (2)	White	76% (146)	South	30% (57)
Man	57% (110)	21-25	8% (16)	Black	6% (11)	Midlands	16% (30)
Non-binary	3% (5)	26-30	18% (34)	Asian	4% (8)	West	19% (37)
		31-35	26% (50)	White, Latine	1.6% (3)	Northeast	24% (46)
		36-40	18% (35)	Other	0.5% (1)	Inland North	7% (14)
		41-45	9% (17)	White, Native Indian	0.5% (1)	North Central	1.6% (3)
		46-50	7% (14)	Black, Indian	0.5% (1)	Declined to answer	2% (4)
		51-55	3% (5)	Declined to answer	3% (6)		
		56-60	6% (11)	Black, Latine	1% (2)		
		>60	4% (7)	Latine	4% (8)		
				White, Asian	1.6% (3)		
				White, Black	0.5% (1)		

Table 4.1. Self-identified demographic information for Study 3 participants.

4.5. Analysis

4.5.1. Old-new judgment task

Two separate logistic mixed effects models were applied for responses to old and new items to determine whether the metalinguistic prime, fall/rise contour, speaker gender, and their interactions were significant predictors of listeners' responses in the old-new task. Old and new items were analyzed separately to inspect the effects of the prime, contour, and speaker gender on listeners' accurate recognition of old items as "old" (hits) and on their inaccurate identification of new items as "old" (false alarms), respectively. Fall/rise contour and male/female speaker gender were contrast-coded identical to Studies 1 and 2. Prime was included as a categorical factor with four levels to assess the effects of different primes and their interactions with other factors (speaker gender and utterance contour) in relation to the baseline condition as a default (no prime). An ANOVA (analysis of the variance) assessed each main effect term's (speaker gender, prime, contour, and their interactions) contribution to the model fit by calculating whether the inclusion of any of these factors lowered the squared sum of residuals from listeners' responses. The interaction between speaker gender and prime did not improve

model fit for either model, and was thus removed from the final models. As stated above, data from 30 participants who misremembered their prime were not included in any of these analyses.

Following Study 1 and Study 2, a post-hoc analysis investigated the effects of perceptual and acoustic factors on listeners' old-new recognition accuracy. Old/new status, Norming 3 contour classification accuracy, and pitch slope differences between falling and rising versions of utterances were included as main effects on listeners' Study 3 old-new recognition accuracy. Speaker, utterance, and participant were included as random intercepts.

Another subsequent analysis compared listeners' old response rates between utterance contours and all four prime conditions (rather than single comparisons to the baseline) via a set of Tukey Honesty Significant Difference (HSD) tests run on the prime factor in each of the models. Separate Tukey HSD tests were conducted on the separate models for old and new items. Utterance contour (fall/rise), metalinguistic prime, and their interaction were loaded onto the Tukey HSD tests for sets of old and new items. These tests compared the following:

1. "Old" response rates for either all rises or all falls between conditions (ex. rises in the baseline condition versus rises in the equal condition)
2. "Old" response rates for rises versus falls within the same prime condition (ex. rises versus falls in the baseline condition)
3. "Old" response rates for utterances of different contours in different prime conditions (ex. falls in the baseline condition versus rises in the equal condition)

Separate Tukey tests for old and new items enabled this analysis to detect differences in listeners' rates of (false) recognition, comparing average "old" response rates for both old items (hits) and new items (false alarms) between the four different prime conditions and two utterance contour types.

A second set of Tukey tests assessed differences in listeners' old response rates for rises between different speaker genders and metalinguistic prime conditions. Similar to the analyses above, responses to old and new items were analyzed in separate Tukey tests. These tests compared the following:

1. "Old" response rates for rises from either all male speakers or all female speakers between conditions (ex. rises from female speakers in the baseline condition versus rises from female speakers in the equal condition)
2. "Old" response rates for rises from male versus female speakers in the same prime condition (ex. rises from female speakers versus male speakers in the baseline condition)
3. "Old" response rates for utterances from different speaker genders in different prime conditions (ex. rises from female speakers in the baseline condition versus rises from male speakers in the equal condition)

This second set of Tukey tests for old and new rises enabled this analysis to detect differences in listeners' rates of (false) recognition for this contour specifically, comparing average hit or false alarm rates when listening to speakers perceived as female versus male and with different metalinguistic priming information.

4.5.2. Cloze task

Listeners' cloze responses were initially hand-checked for typographical errors, then scored as 0 or 1 for response accuracy following the previously described accuracy schema (Chapters 2 and 3). Listeners' accuracy scores were submitted to a mixed effects logistic regression, predicting accuracy by fall/rise utterance contour, prime condition, speaker gender, and the interactions between prime with contour and with speaker gender. A mixed effects linear regression was fit to

predict listeners' old-new accuracy by their cloze accuracy scores to test the correlation between accuracy scores of these two tasks. Listeners' average old-new accuracy combined old and new utterances and was classified as a match between the utterance status and listener response. Participant was included as a random intercept.

4.6. Predictions

4.6.1. Old-new judgment task

My predictions and analysis for Study 3 attended to listeners' "old" response rates across each of the four different meta-linguistic primes. If a metalinguistic prime affected the ways in which listeners remember falling and rising utterances, then I expected their hits and false alarms to be influenced by this information. If a metalinguistic prime can induce attention to rising utterances, thus leading to better recognition, then listeners were expected to exhibit more hits ("old" response to old items) for rises than falls in prime conditions versus the baseline condition. If the metalinguistic prime overpowered the encoding of real linguistic utterances such that listeners falsely remembered those utterances mirroring their prime, then rises were expected to garner more false alarms ("old" responses to new items) in prime conditions than in the baseline condition without a prime. Listeners' rates of (false) recognition of falling utterances were expected to be consistently lower than rising utterances in those primed conditions. This would have been illustrated by significantly higher "old" response rates for rises in contrast to falls in prime conditions versus the baseline condition (no prime).

The specific content of a listeners' prime was predicted to further influence their hits and false alarms. If the metalinguistic prime was successful at inducing a gender-related bias about uptalk, then I predicted that listeners' responses to rising utterances would be mediated by this

information. A gender-related bias may have boosted listeners' memories of rising utterances when produced by a speaker whose gender was referenced in their prime. Specifically, listeners who received the male-biased prime were predicted to exhibit more "old" responses for rises when listening to male speakers, in contrast to rises by female speakers. To support this prediction, a three-way interaction between metalinguistic prime, contour, and speaker gender would yield higher "old" response rates for rises in contrast to falls, from male speakers in the male-biased prime condition.

Listeners who received the female-biased prime were also predicted to exhibit a gender bias toward female speakers' rises. I expected listeners in this condition to exhibit greater "old" response rates for female speakers' rises than those participants in the baseline condition. Since broader metalinguistic discourse on uptalk in the U.S. lends more support for the female-biased prime (Grose 2015), I also predicted that in relation to the baseline condition, the female-biased prime, in contrast to the male-biased prime, would induce more hits and false alarms for female speakers' rises. These effects would be illustrated by a significant interaction between speaker gender and utterance contour in the female-biased prime.

4.6.2. Cloze task predictions

Listeners' responses in the cloze task, as suggested from the previous experiments in this dissertation, have a complex relationship to their performance in the old-new task. Following Studies 1 and 2, cloze accuracy was not predicted to exhibit strong correlations with old-new accuracy or be significantly different between falls and rises. However, the presence of a metalinguistic prime about uptalk may have boosted listeners' accuracy in remembering cloze items previously produced with rising intonation. Listeners' cloze accuracy was therefore predicted to be influenced by the interaction between utterance contour and metalinguistic prime.

4.7. Results

4.7.1. Old-new judgment task

4.7.1.1. Hits ("old" responses to old items)

Listeners' responses to old items across all four conditions and their main effects are described in Table 4.2. below.

Predictor	Estimate	Std. Error	Z-value	P-value
(Intercept)	0.917	0.135	6.770	<0.0001***
Fall/rise contour (= rise)	-0.149	0.244	-0.611	0.541
Speaker gender (= man)	-0.075	0.200	-0.377	0.707
Equal prime	0.027	0.157	0.175	0.861
Female-biased prime	-0.209	0.136	-1.542	0.123
Male-biased prime	-0.263	0.143	-1.836	0.066
Contour * speaker gender	0.287	0.358	0.801	0.423
Contour * equal prime	-0.112	0.291	-0.385	0.700
Contour * female-biased prime	0.621	0.251	2.477	0.013 *
Contour * male-biased prime	-0.106	0.265	-0.401	0.688
Contour * speaker gender * equal prime	-0.396	0.582	-0.681	0.496
Contour * speaker gender * female-biased prime	-1.159	0.503	-2.306	0.021 *
Contour * speaker gender * male-biased prime	-0.043	0.530	-0.082	0.935
Norming 3 contour accuracy	-0.445	0.543	-0.819	0.413

*Table 4.2. Mixed effects logistic regression summary Study 3 listeners' "old" responses to old items. N=2,156; * = $p < 0.05$, ** = $p < 0.005$, *** = $p < 0.0005$.*

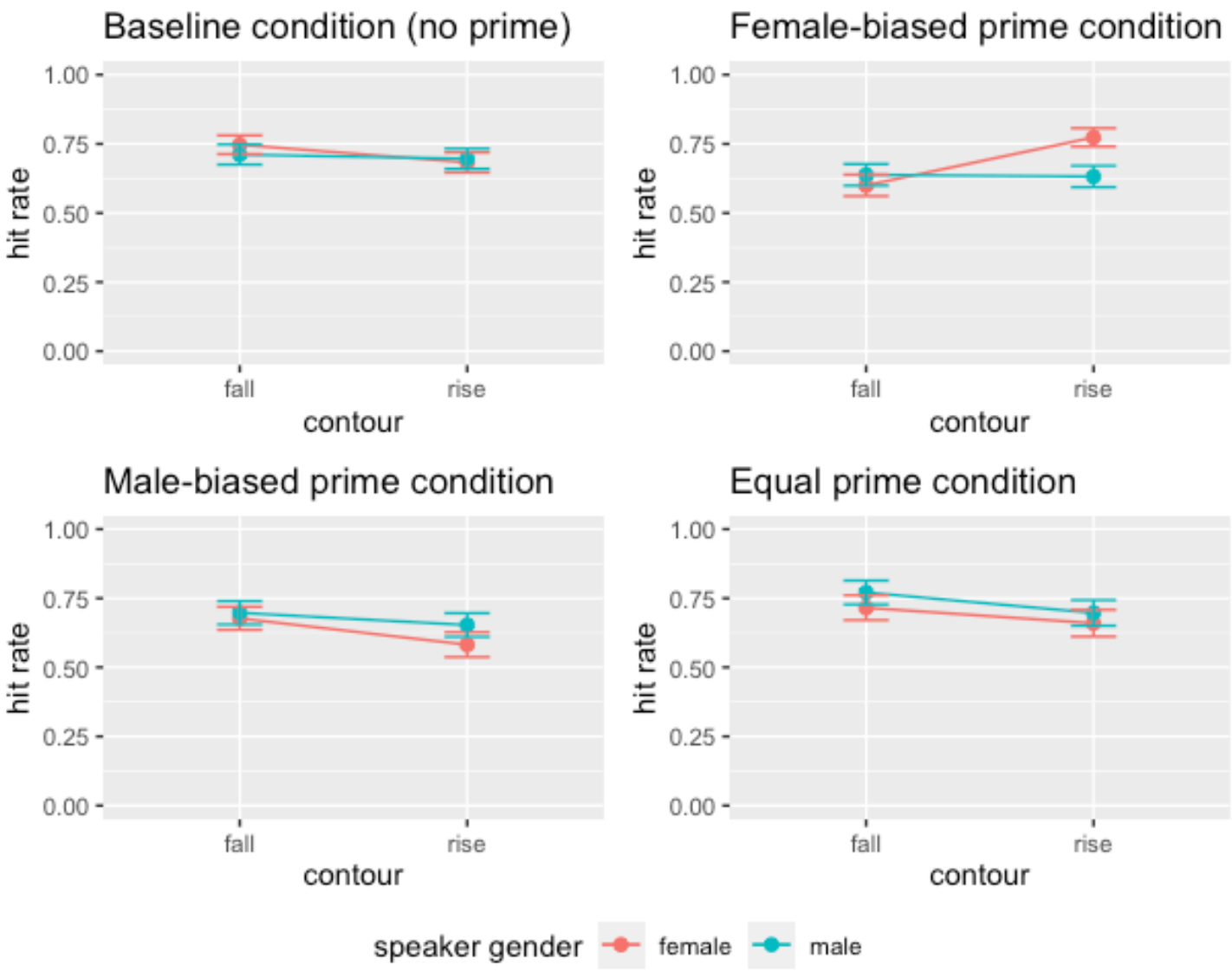

Figure 4.1. Listeners' "old" responses for old item across four conditions, colored by speaker gender.

Listeners' hits for falls and rises were predicted to vary across prime conditions. Specifically, listeners who heard a metalinguistic prime about uptalk were expected to exhibit more hits for rises than falls in contrast to those participants without a prime in the baseline condition. Results illustrate that the only group of listeners who exhibited this effect were those in the female-biased prime condition (Figure 4.1, top right). Participants who were presented with the statement that women were most likely to use uptalk were more likely to accurately classify rises as "old" at greater rates than their accurate classification of falls than participants in other prime

conditions were (Table 4.2). This effect provides partial support for my predictions that listeners' responses would reinforce a pre-existing stereotype: rather than any metalinguistic prime leading to greater hits for rises than falls, participants exhibited greater hit rates for rises only when they were presented with the female-biased prime.

Results from the Tukey test comparisons for hit rate differences the two contour types, compared among the four prime conditions, illustrated (Appendix G, Table 1) that listeners' hit rate for falls in the baseline condition was significantly greater than their hit rate for falls in the female-biased condition (Diff = 0.111, p = 0.05). This corroborates results from the logistic regression analysis of listeners' hit rates: When listeners received the female-biased prime about women's use of uptalk, they less frequently recognized female speakers' falls than listeners who were presented with no prime. This prime and contour effect suggests that listeners' recognition of rises was greater when they received information prompting women's use of uptalk versus when they received other metalinguistic information (or none), and listeners' recognition of falls was weaker with the provision of this information in contrast to when listeners received no prime in the baseline condition.

Furthermore, I predicted that the gender bias of the prime would modulate listeners' hit rates for rises when listening to speakers of different perceived genders. Specifically, I predicted that a prime highlighting women's use of uptalk would lead to greater hits for female speakers' rises versus falls, and a prime highlighting men's use of uptalk would lead to greater hits for male speakers' rises versus falls. Listeners' responses indicated a significant three-way interaction between speaker gender, contour, and metalinguistic prime condition. Results indicated that listeners were more likely to recognize rises more than falls for speakers perceived as women versus as men, but only when they were exposed to the female-biased metalinguistic

prime (Table 4.2). In this way, when listeners were told that women use uptalk more than other speakers, they exhibited more hits for rises than for falls when listening to a speaker perceived as a woman versus as a man. The interaction between speaker gender and contour did not occur with any other metalinguistic prime condition other than the female-biased prime. In other words, only listeners who were exposed to the female-biased prime showed greater hit rates for recognizing rises versus falls when listening to a speaker whose perceived gender matched the metalinguistic prime.

Results from the Tukey test for hit rates to rises between different speaker genders and prime conditions revealed one combination of speaker gender and prime that yielded a significant difference in listeners' responses (Appendix G, Table 2). Listeners exhibited more hits for rises from female speakers in the female-biased condition than in the male-biased condition (Diff = -0.192, $p = 0.02$). This result provided confirmation of the effect of the metalinguistic prime in the logistic regression: When primed to expect women's use of uptalk, listeners' recognition for women's rises was significantly higher than when prompted toward men's use of uptalk. However, I also predicted that the female-biased prime would be stronger in inducing a gender bias than the male-biased prime, such that listeners with the female-biased prime would exhibit greater hit rates for rises for female speakers with the female-biased prime than for male speakers with the male-biased prime. The difference in hit rate between rises from male speakers in the male-biased condition and rises from female speakers in the female-biased condition was not statistically significant ($p = 0.38$). All together, results from listeners' "old" responses to old items (hit rate) indicated that metalinguistic information stating that women were more likely to use uptalk than men successfully primed listeners to remember speakers' rises, especially when those speakers were perceived as women. Other priming statements about uptalk, as well as the

absence of any metalinguistic information, did not deliver the same effect in leading listeners to recognize rises in general or in any other gender-biased direction.

4.7.1.2. False alarms ("old" responses to new items)

Listeners' old-new responses to new items across all four conditions and their main effects are described in the Table 4.3. below.

Predictor	Estimate	Error	Z-value	P-value
(Intercept)	0.238	0.102	2.33	0.020 *
Fall/rise contour (= rise)	-0.399	0.193	-2.066	0.039 *
Speaker gender (= fall)	0.013	0.077	0.165	0.869
Equal prime	-0.076	0.113	-0.671	0.502
Female-biased prime	-0.075	0.113	-0.663	0.507
Male-biased prime	-0.004	0.113	-0.035	0.972
Contour * speaker gender	0.014	0.148	0.095	0.924
Contour * equal prime	0.023	0.208	0.11	0.913
Contour * female-biased prime	0.851	0.208	4.083	<0.0001 ***
Contour * male-biased prime	0.402	0.209	1.924	0.054
Norming 3 contour accuracy	0.559	0.399	1.399	0.161

*Table 4.3. Mixed effects logistic regression summary Study 3 listeners' "old" responses to new items. N= 2,156; * = $p < 0.05$, ** = $p < 0.005$, *** = $p < 0.0005$.*

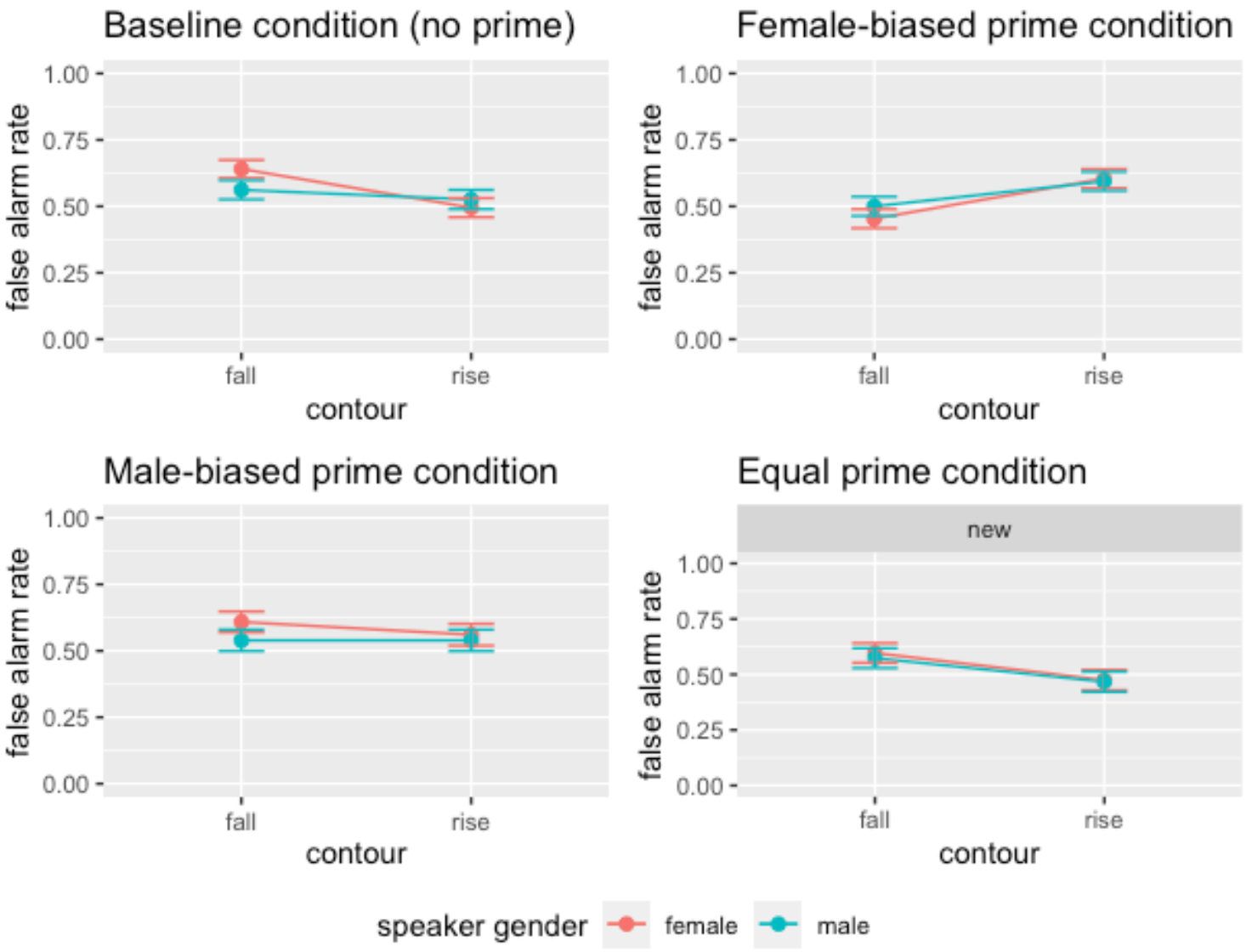

Figure 4.2. Listeners' false alarms ("old" responses for new items) across four conditions, colored by speaker gender.

For listeners' responses to new items, I first predicted that listeners who received any metalinguistic prime would exhibit greater false alarms ("old" responses to new items) for rises than those participants in the baseline condition who did not receive a prime with an explicit mention of uptalk. The presence of different metalinguistic primes did not significantly lead to more false alarms for rises than when listeners did not receive any metalinguistic information (Table 4.3). Results from the mixed effects logistic regression model did indicate fall/rise

contour is an independently significant main effect: Listeners on average false alarmed (inaccurately classified new item as "old") for falls more than for rises (Table 4.3). One prime condition significantly modulated this effect, indicated by an interaction between contour and prime. The female-biased prime condition, in which listeners were told that women were more likely to use uptalk than speakers of other genders, garnered more false alarms for rises than for falls (Table 4.2) in contrast to participants who received no metalinguistic information (baseline prime condition) or a different metalinguistic priming statement. This effect suggested partial support for my predictions: in contrast to other primes or to no prime at all, a metalinguistic prime highlighting women's use of uptalk garnered more false memories of rising utterances, regardless of speaker gender.

Regarding results from the Tukey test for listeners' false alarms, three combinations of prime and contour showed significant differences within the same contour type or prime (Appendix G, Table 3). Additionally, responses significantly differed between those to equal prime responses to rises and those to baseline responses to falls (Diff = -0.131, p = 0.027), though this is somewhat difficult to interpret due to the interaction of two independent variables in this comparison, so no predictions were made for this contrast. Listeners in the female-biased prime condition exhibited a significantly higher false alarm rate for rises than falls (Diff = 0.122; p = 0.015). Rises also garnered more false alarms for listeners in the female-biased prime than those in the equal prime (Diff=0.128, p=0.037). Finally, listeners' false alarms to falls were significantly lower in female-biased prime condition than in the baseline (no prime) condition (Diff = -0.125, p = 0.0098). These Tukey test results provide support for the interpretations of the results from the mixed effects logistic regression: When told that women used more uptalk than others, listeners were more likely to falsely remember rises than falls, and false alarms for

rises in this condition were greater than false alarm rates when they were told that both women and men use uptalk at the same rate in the equal prime condition.

I predicted that the male- and female-biased primes would also boost listeners' false alarms for rises from male and female speakers, respectively. The interaction between speaker gender, contour, and prime was not found to be a significant predictor of listeners' false alarms. Unlike listeners' hit rates to falls and rises, their false alarms to these contours were not modulated by the perceived gender of the speaker. Furthermore, results from the Tukey test found that listeners' false alarms to rises from speakers of different perceived genders were not significantly different from one another in any prime condition (Appendix G, Table 4). These results do not provide support for my predictions regarding the effects of gender-biased primes: The metalinguistic prime did not significantly interact with the perceived gender of the speaker, failing to induce greater false memories of rises when listening to a speaker whose perceived gender was highlighted or omitted from their prime.

4.7.2. Post-hoc analysis of perceptual and acoustic factors and old-new recognition

Predictor	Estimate	Std. Error	Z-value	P-value
Intercept	1.528	0.701	2.182	0.029 *
Old/new status (=old)	0.488	0.189	2.581	0.01 *
Norming 3 contour accuracy	0.855	0.704	1.214	0.225
Rise-fall slope difference	-0.0005	0.002	-0.316	0.752

*Table 4.4. Mixed effects logistic regression summary for post-hoc analysis of Study 3 listeners' old-new recognition. N=1,536; * = p < 0.05, ** = p < 0.005, *** = p < 0.0005.*

In post-hoc analyses of Block 1 old-new judgment performance, neither the factors for Norming 3 contour accuracy nor for pitch slope difference was to be a significant predictor of listeners' accuracy in their old-new judgments. Regardless of how accurately Norming Study 3 listeners classified utterances as a fall or rise, or how contrastive the pitch slopes were between falling and rising versions of utterances, listeners' accuracy was predicted by the old/new status of the

utterance. Listeners were more likely to accurately classify old utterances as "old" over "new" utterances as "new" (Table 4.4).

4.7.3. Cloze task

Listeners' cloze accuracy and their main effects (contour, speaker gender, prime, and their interactions) are described in Table 4.5 below.

Predictor	Estimate	Std. Error	Z-value	P-value
(Intercept)	-0.666	0.408	-1.631	0.103
Fall/rise contour (=rise)	0.993	0.633	1.568	0.117
Equal prime	-0.024	0.118	-0.200	0.841
Speaker gender	-0.088	0.166	-0.528	0.598
Female-biased prime	0.013	0.118	0.111	0.912
Male-biased prime	-0.170	0.119	-1.427	0.154
Contour * equal prime	0.072	0.236	0.302	0.762
Contour * female-biased prime	-0.219	0.236	-0.930	0.353
Contour * male-biased prime	-0.176	0.239	-0.739	0.460
Equal prime * speaker gender	-0.09	0.236	-0.365	0.715
Female-biased prime * speaker gender	-0.181	0.235	-0.773	0.440
Male-biased prime * speaker gender	0.07	0.28	0.325	0.745

*Table 4.5. Mixed effects logistic regression summary of Study 3 listeners' cloze accuracy. N=3,100; * = $p < 0.05$, ** = $p < 0.005$, *** = $p < 0.0005$.*

Listeners' cloze accuracy was predicted to be influenced by the contour, perceived speaker gender, and prime that they either did or did not receive in their experimental condition. None of these factors nor their interaction was shown to significantly modify listeners' accuracy in the cloze task. These results did not support my predictions for the effects of prime and utterance contour on listeners' cloze performance. Furthermore, listeners' cloze accuracy was not found to be a significant predictor of their accuracy in the old-new judgment task. Following Studies 1 and 2, analyses in Study 3 indicate no significant relationship between listeners' performance between the two memory tasks.

4.8. Discussion

4.8.1. Metalinguistic prime effects

The aim of Study 3 was to assess whether metalinguistic information about uptalk mediated listeners' false memories of falls and rises from speakers perceived as different genders. I posited that the presentation of a metalinguistic statement about gender in relation to the use of uptalk would influence the ways in which listeners recognized items with different intonational contours. Specifically, if a metalinguistic prime about uptalk can bias listeners' expectations of certain speakers' rise productions, then "old" responses for rises from primed listeners were predicted to be greater than "old" responses from listeners without a prime. Listeners' responses revealed that the effect of prime was only significantly different from no prime in one condition, the female-biased prime condition. When listeners were told that women used uptalk more than speakers of other genders, they exhibited greater "old" responses to both old items (hits) and new items (false alarms) for rises than for falls. The contour effect in this prime condition reversed in the other three conditions, including the baseline condition: listeners exhibited greater "old" responses for falls more than rises.

Interpreting these results, information regarding *women's* uptalk in particular appeared to most prominently boost listeners' expectations of having heard rising contours, in contrast to information about men's uptalk or equal use across all speakers. The significant difference with which listeners (falsely) recognized rises over falls in the female-biased prime condition compared to all other conditions suggests that the effect of the metalinguistic prime could be influenced by listeners' existing metalinguistic knowledge. A prime that reaffirmed existing stereotypical associations between women's voices and uptalk was able to enhance listeners' recognition and induce false memories of rising utterances. When listeners received statements

containing less stereotypical information about uptalk, as in the equal prime or male-biased prime condition, they behaved similarly to those participants who received no prime at all, with greater hits and false alarms for falls than rises. In this way, Study 3 listeners' awareness of stereotypical metalinguistic information may have modulated the ways in which they remembered intonational productions from their speaker. The greater rates at which listeners (falsely) recognized rises in the female-biased prime condition suggested that listeners' awareness of the socially meaningful link between uptalk and women's voices provided greater weight to the encoding of rising utterances. Supporting the conclusions put forth by Sumner et al. (2014), the social information embedded in discourses around this feature appeared to garner more robust activation of listeners' episodic memory of rises for women, yielding greater hits and false memories of these productions.

I also predicted that, following metalinguistic discourses that focuses on women's usage of uptalk (e.g. Davis 2010), the female-biased prime would be stronger than the male-biased prime in inducing a rise bias with a speaker of the referenced gender. Specifically, listeners' "old" responses for female speakers' rises in the female-biased prime condition were predicted to be greater than "old" responses for male speakers' rises in the male-biased prime condition. The female-biased prime, unlike the male-biased prime, effectively induced a gender bias such that listeners provided more hits for rises over falls for speakers perceived as women, with no contour effect for speakers perceived as men.

These results lend further support for the influence of metalinguistic stereotypes in listeners' recognition of falling and rising utterances. When listeners were prompted to expect uptalk from speakers perceived as women, they were more accurate in recognizing rises for female speakers than male speakers. Listeners who received the male biased prime did not

appear to enjoy the same advantage of greater memory for male speakers' rising productions. Interpreting these results, the salient stereotypical link that was present in the female-biased prime may have attributed greater social weight with which listeners encoded episodes of rising productions from speakers perceived as women, over falling productions from those speakers or rising productions from male speakers. In this way, the social importance of the female-biased prime in relation to the other primes (or lack of any information in the baseline) facilitated listeners' accurate recognition of rising utterances that they heard before from female speakers.

However, listeners' false alarms did not exhibit a speaker gender effect. Regardless of which gendered prime they heard, listeners' false alarms for rises did not significantly differ between male and female speakers. In other words, the prime did not successfully induce more false memories of rises from the targeted speaker gender. Rather, the female-biased prime yielded greater false memories for all rises in contrast to falls for both speakers perceived as male and female. The female-biased prime appeared to successfully bias listeners toward accurately remembering rises from female speakers over male speakers and female speakers' falls, but not effective enough to induce false memories for rises exclusively from female speakers.

Listeners' false alarms across different prime conditions and the baseline provide some complications to results from Studies 1 and 2. A metalinguistic prime about uptalk was most effective in boosting listeners' false memories of rises when the prime fixated on women's voices in contrast to other speakers. In contrast, a prime solely about uptalk use, as in the equal prime, failed to induce a bias toward rises over falls: Only that prime which specifically highlighted women's voices, and corroborated popular and academic stances (Warren 2016; Linneman 2013), boosted listeners' false memories of this intonational contour. These effects,

while different than what was predicted, provide evidence for the prevalence of a cognitive link between gendered expectations and memories of uptalk. Overall, results from Study 3 suggest a prime about the use of socially meaningful linguistic features can influence listeners' recognition of intonational contours, and thereby potentially influence the ways in which speech is encoded and recalled, but this effect appears to vary based on the specific social content of the metalinguistic prime in ways that uphold pre-existing social stereotypes.

Regarding post-hoc analyses of listeners' old-new recognition accuracy, accuracy scores were consistently lower for those utterances that were most accurately classified in Norming Study 3. This trend exhibited the opposite direction expected by my predictions: Rather than boosting listeners' accuracy in distinguishing between old and new intonation, utterances perceived as prototypical falls or rises led to less accurate recognition rates. While this interpretation is speculative and requires further investigation, utterances that may have exhibited less distinctive cues of these intonational categories of falling or rising intonation (exhibited by lower accuracy of listeners' contour classification in Norming Study 3) could have been attended to more closely in the task of recognizing old versus new intonation. This in turn may have yielded more accurate responses for these utterances versus those that were more consistently classified as their respective contour.

4.8.2. Listeners' recall of and agreement with prime

Listeners' existing social lenses mediate how they interpret metalinguistic evaluations of others and themselves (Jaworski et al. 2012; Aslan & Vásquez 2018). Their understandings of themselves and others in their social world – guided by both direct and indirect experiences as well as by socially learned connections – may interact with an induced statement about generalized sociolinguistic patterns. To consider listeners' own stances and expectations around

who is most likely to use uptalk, for those who received a metalinguistic prime, a post-hoc analysis was conducted on listeners' qualitative commentary to assess whether they remembered and agreed with their prime. As described above, listeners first indicated which gender of speakers, according to their prime statement, was most likely to use uptalk. Listeners then explained whether or not they agreed with their prime, in an open-response format. Listeners who misremembered their prime were included in this post-hoc analysis in order to retain their data regarding the social associations each participant held between uptalk and perceived speaker gender. Listeners' responses to these questions were not submitted to statistical tests or included in any models of old-new judgment or cloze performance. Rather, responses were qualitatively assessed for any differences among prime conditions. The distribution of listeners' qualitative responses regarding their prime is discussed below and illustrated in Figures 4.3-5 below.

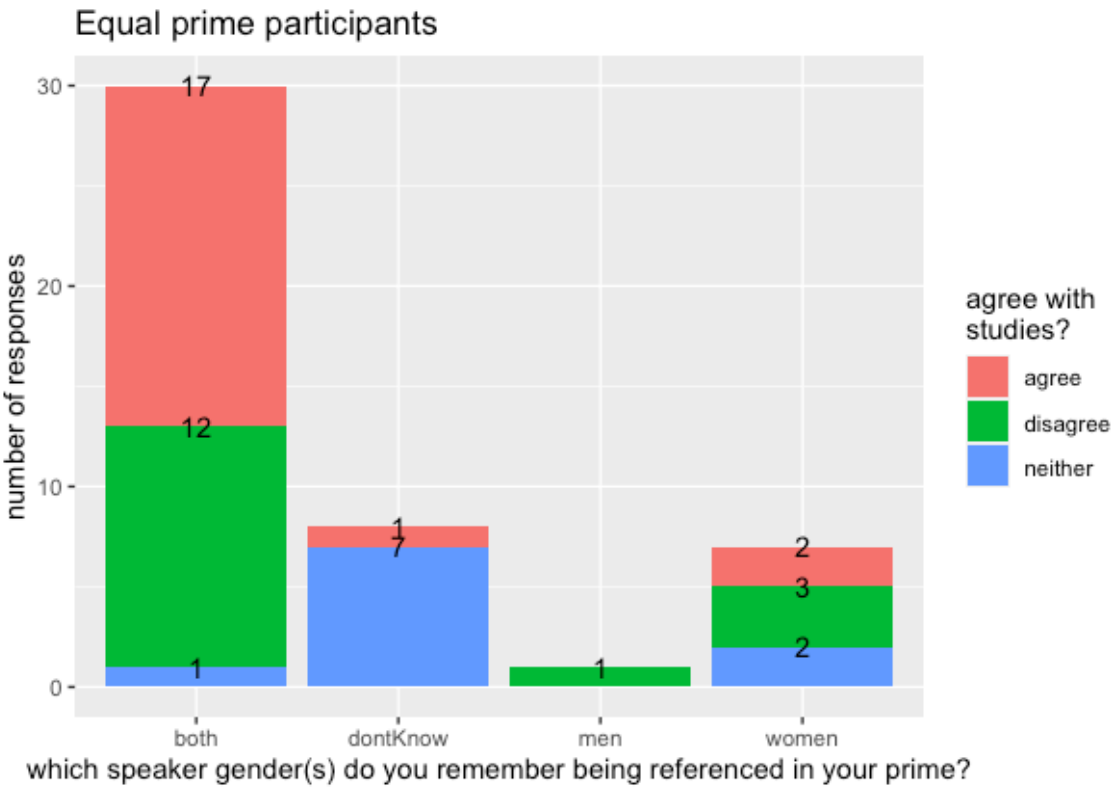

Figure 4.3. Counts of responses for listeners' recollection of the speaker gender specified in the equal prime, colored by their agreement with the prime.

When presented with the equal prime which stated that "recent studies show that men and women are equally likely to use uptalk" (Figure 4.3), most participants accurately indicated that they were told that "both" men and women use uptalk. However, almost half of those participants disagreed with this prime. Other participants in the equal prime condition indicated that they did not remember which gender their prime referenced or falsely recalled that the prime referenced a different gendered group. Out of the 12 participants in the equal prime condition who disagreed with their prime, 10 of them explicitly argued in the open-response question that women were more likely than others to use uptalk, often citing personal beliefs and experiences:

1. 51-year-old white man from the West: "I feel like women do it more. It makes me think of 'Valley Girls' and the way they speak. I think men speak more monotone in geral [sic]."
2. 57-year-old white man from the Northeast: "[No, because] it's been my experience that it's mostly women who use uptalk."
3. 28-year-old white woman from the Northeast: "I definitely hear uptalk used more in women."
4. 49-year-old white woman from the South: "I find more women using uptalk."

The other two participants who disagreed with the equal prime indicated they would need more evidence and experience to agree with this statement.

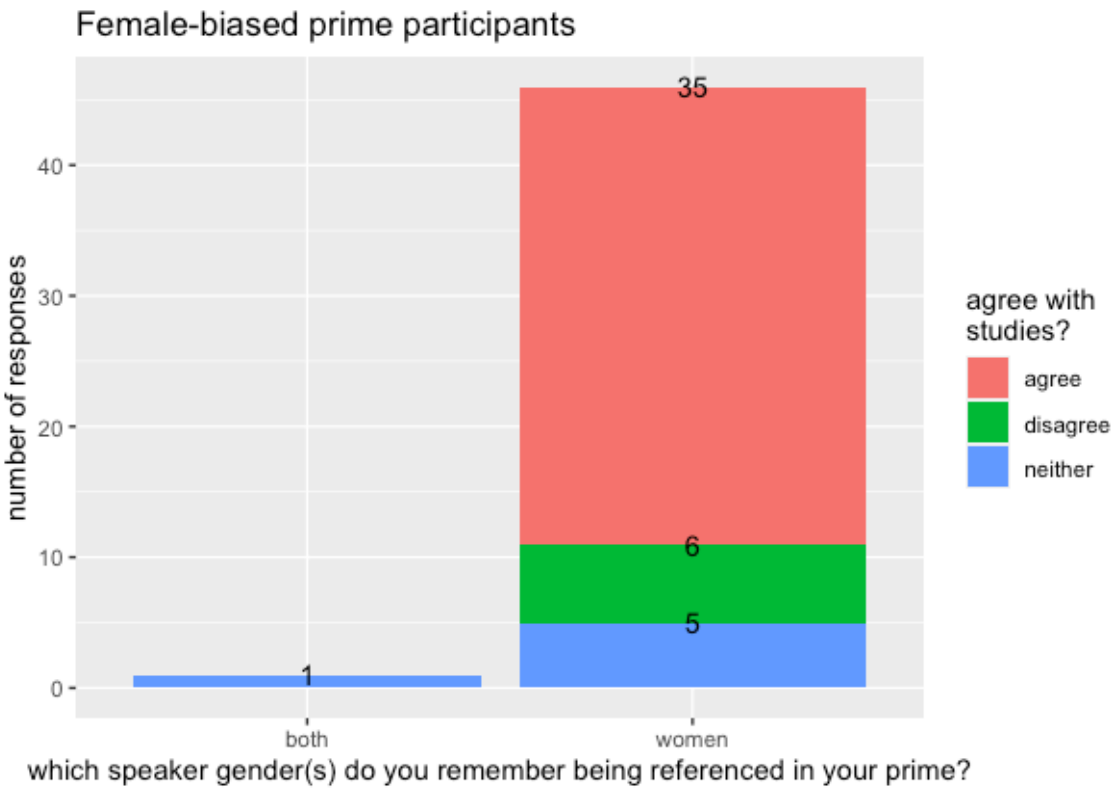

Figure 4.4. Counts of responses for listeners' recollection of the speaker gender specified in the female-biased prime, colored by their agreement with the prime.

Listeners in the female-biased prime (Figure 4.3) condition almost categorically exhibited no confusion about the content of their prime: No participant indicated "I don't know" when responding to which gender group was referenced in their prime, and only one participant in this condition incorrectly recalled that their prime stated that "both" gendered groups use uptalk. Nearly 75% of those participants who accurately recalled their female-biased prime agreed with its statement in their explanation. In their responses, 18 participants out of the 35 who agreed with the female-biased prime explicitly referenced their own lived experiences in which they remember women using more uptalk than other gendered speakers. A number of participants underscored their agreement by overtly naturalizing the link between femininity and uptalk, "making sense" of this information:

1. 29-year-old white man from the Inland North: “Yes, it makes sense [women using more uptalk than others]”
2. 41-year-old white man from the South: “I agree, because it [uptalk] sounds feminine.”
3. 58-year-old white man from the Inland North: “Yes, it sounds stranger coming from a man and since that is the case, I’m guessing I hear it more from women and never thought about it.”
4. 34-year-old white woman from the South: “yes because women naturally have an uptalk most of the time.”
5. 23-year-old Black woman from the South: “I agree because when I notice uptalk, it is usually from a woman. I find it to typically be a very feminine way of speaking.”
6. 23-yeard-old Black/Native Indian woman from the Northeast: “I agree because it seems to be a feminine behavior.”
7. 24-year-old white woman, region not defined: “It makes sense to me that women would be more likely to use uptalk than men.”
8. 42-year-old Black woman, region not defined: “Yes, women naturally do this because a lot are mothers and they use the tone with their children.”

Relatively few participants (two out of 35) attributed their agreement with the female-biased prime to the prevalence of existing social pressures and structures. These two participants both self-identified as women and provided further explanations regarding these notions:

1. 29-year-old Black woman from the Northeast: “I do [agree]. Women are more socialized than men to be more appealing in various ways, including the way they talk.

2. 24-year-old white woman from the South: “I do, because many women are trained/influenced to manipulate their voices to be higher due to gender stereotypes.”

Seven participants disagreed with the female-biased prime, the lowest number of participants who correctly identified their prime and disagreed with its statement. Five participants who disagreed with the female-biased prime argued that other speakers were just as likely to use uptalk as female speakers:

1. 38-year-old white man from the West: “In the age of ‘influencer’ and ‘youtuber’, no. They all do it [use uptalk].”
2. 26-year-old non-binary participant from the South: “To me it seems like I hear men doing that [uptalk] just as often.”
3. 44-year-old white woman from the West: “I don’t have enough insight to disagree, but most of the auditory examples, male and female, tended to use the uptalk.”
4. 60-year-old white woman from the West: [I] Believe that it [uptalk] is a habit that an insecure speaker uses, regardless of their gender.
5. 32-year-old white man from the Northeast: “I think uptalk is common in both genders for varying reasons.”

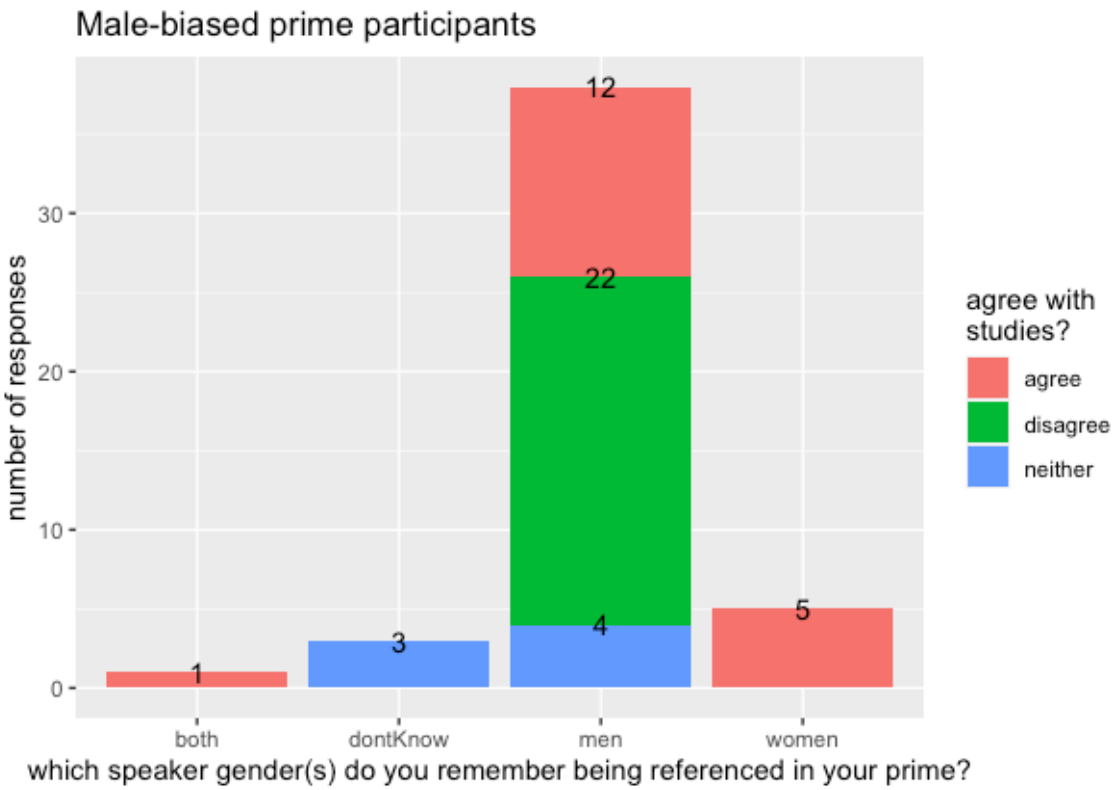

Figure 4.5. Counts of responses for listeners' recollection of the speaker gender specified in the male-biased prime, colored by their agreement with the prime.

When receiving the prime stating men use more uptalk than other genders (Figure 4.5), most participants (38 out of 48 in this condition) accurately recalled this prime. However, more than half of these participants disagreed with its statement, in contrast to listeners' overwhelming agreement with the female-biased prime. The majority of listeners who disagreed with their male-biased prime (14 participants out of the 22 who disagreed with this prime) argued that they believed women, in contrast to men, use more uptalk, referencing their personal experiences and expectations. However, the remaining proportion of participants who disagreed with their male-biased prime (eight in total) did so because of their belief that speakers of all genders, including men and women, use uptalk at similar rates:

5. 34-year-old Asian man from the West: "No. Regardless of their gender, I've heard a lot of people using uptalk."

6. 25-year-old white man from the Inland North: “I disagree with this study’s findings because I witness men and women using uptalk about evenly in my own experiences in life.”
7. 27-year-old white man from the South: “I don’t agree because I think both men and women use uptalk, sometimes unconsciously.”

A total of five participants out of 47 participants in the male-biased prime condition inaccurately remembered the prime as stating that women use more uptalk than other genders, and they all indicated agreement with this statement in their comments.

Listeners’ commentary about men’s and women’s uptalk, paired with trends found in Figure 4.2, align with persistent mainstream discourses about uptalk and who uses this feature. Listeners, when told that women use uptalk more than other speakers, exhibited little to no confusion in recalling their prime. The majority of these female-biased primed participants signaled agreement with this prime, mirroring larger-scaled metalinguistic commentary that points to women’s greater use of uptalk and circulates negative social stereotypes about these voices (e.g. Davis 2010). The proportion of listeners who misremembered their male-biased prime as referencing women’s use of uptalk (11%) particularly speaks to the prevalence of the link between female gender expression and uptalk. Listeners’ disagreement with their primes, however, complicate this trend. Primed participants did not categorically agree that women used more uptalk than other gendered speakers; rather, the sample of participants provided wide-ranging arguments, citing lived experiences and personal beliefs, against this gendered stereotype about women’s voices. However, a stance against this metalinguistic statement was relatively rare in the participant sample, and it was even outnumbered by participants who mis-identified their prime as one mentioning women’s uptalk.

Listeners' responses to these questions also provide support for interpretations of results from the old-new judgment task. As the female-biased prime was the strongest out of all conditions to induce a rise bias in listeners' hits and false alarms, listeners' recall of this prime illustrates that this information about women's use of uptalk was especially memorable for listeners. As stated in Section 4.2.2, The proportion of listeners who falsely recalled their prime were excluded from the analyses of their old-new recognition, but their performance in the experiment nevertheless indicated an influence of stereotypical expectations on memory of speech and metalinguistic information. In particular, those participants who misremembered their male-biased or equal prime as mentioning women's voices (12 participants out of 196 in the sample) further illustrate that individuals are more likely to recall information in way that aligns with their existing attitudes or beliefs – namely, their expectations of which gender of speakers uses uptalk most frequently. In this way, both groups of listeners who correctly or incorrectly recalled their prime exhibited a bias toward remembering linguistic content (both the metalinguistic prime and old-new tokens) that aligned with the social stereotype linkage of women's voices with uptalk.

However, it is noteworthy that as listeners' open responses indicated a cognitive and attitudinal preference for remembering and agreeing with the female-biased prime, it was also this prime condition that led to greater "old" response rates to rises in the old-new judgment task. This pattern provides evidence for congeniality effects for stereotypical links between uptalk and women: Listeners, despite being presented with one gendered association, falsely remembered a different statement that was substantiated by pervasive metalinguistic commentary on women's use of uptalk (Warren 2016). This suggests that listeners' interpretation of metalinguistic information is constrained by their understanding of existing stereotypes about linguistic

features, and this in turn mediates how such information influences how they remember old and new intonational productions.

Altogether, listeners' responses in the memory tasks and metalinguistic prime questions in Study 3 indicate that messaging about women's use of uptalk, in contrast to men's, significantly heightened the degree to which listeners accurately recognized rises from female speakers over male speakers, leading to more "old" responses in both old and new items to rises than falls from speakers of both perceived genders. Listeners were most likely to accurately recall and agree with the metalinguistic statement that echoed pervasive metalinguistic links between women's voices and uptalk. In order to more fully investigate the parallels between participants' commentary and larger-scaled metalinguistic discourse about uptalk, we must await future analyses to investigate the degree to which certain listener factors (age, gender, racialized identity, geographic origin) influence listeners' endorsements of different metalinguistic stances about gendered speakers and uptalk.

5. Exploratory factors in listeners' sociolinguistic backgrounds and old-new recognition

5.1. Introduction

Research in social psychology has revealed that individuals engage with stereotypes that encompass pre-existing expectations and attitudes about social groups (Vinacke 1957; Augoustinos & Walker 1998; Macrae & Bodenhausen 2001; Tajfel 2010). Attitudes and stereotype endorsements are not only guided by individuals' own personal experiences; they also can be constructed via passively learned ideological reasoning (De Houwer, Thomas, & Baeyens, 2001; Hargreaves & Tiggemann 2003). Listeners' attitudes have also been shown to affect evaluations of speakers (Campbell-Kibler 2010) and their own sociolinguistic perceptions (Levon 2014).

For example, Levon's (2014) work in sociolinguistic perception showed that attitudes toward gender and sexuality corresponded to social evaluations of speech. Further, Loudermilk (2015) illustrated implicit associations between linguistic attitudes about the sociolinguistic variable (ING) and listeners' brain activity in speech processing. Specifically, listeners who exhibited less prominent stereotypical associations between *–ing/–in* variants and intelligence/education attributes, indicated by low scores in an Implicit Associations Test (IAT), differed in electroencephalogram (EEG) brain-activity from higher-scoring IAT participants who had shown more prominent implicit associations between ING variants and intelligence or education (Loudermilk 2015). Low-scoring IAT participants' linguistic attitudes appeared to exhibit greater N400 negativities when processing *–ing/–in* word variants that violated expectancies based in dialectic variation (ex. The velar variant *–ing* uttered by a Southern-identified speaker and the alveolar variant *–in* spoken by a Californian-identified speaker). In this way, the prominence of listeners' ideologically bound expectations about sociolinguistic variants

altered the fine-grained neural and cognitive activity of perceiving productions of these variants couched in lexical items.

Listeners' explicit memory of speech may also be subject to underlying linguistic attitudes, listeners' own sociolinguistic positionality, and related social expectations toward their speaker. As discussed in Chapter 1, individuals' social positioning, conceptualized as the ways in which they position themselves in relation to underlying systems of societal power, prestige, and privilege (Babel 2016; Bialecki 2021), guide their sociolinguistic practice (Labov 1972; Irvine and Gal 2000). Individuals use and perceive language through the crafted lenses of their own social worlds (Irvine & Gal 2000). As sociolinguistic work already posits that components of a listener's identity affects speech perception (e.g. Ikeno & Hansen 2007; Wagner & Hesson 2014), the investigation of the interaction among different aspects of a listener's social positioning would provide crucial insight into how an individual's linguistic cognition is essentially shaped by their previous experiences of perceiving others' voices as well as being judged themselves.

Analyses in this chapter aimed to study listeners' sociolinguistic positionality not only as a product of their lived experiences and attitudes, but also as a factor in the ways in which they store socially meaningful speech features in memory. I analyzed four aspects of listener positionality – their personal experiences regarding their own voice, attitudes about others' voices, their gender, and their age – as potential influences on listener memory. The analyses of listeners' performance in the old-new recognition task were an opportune space for this exploration because it has collected a sample of English-speaking listeners from a wide range of sociolinguistic and demographic backgrounds as remote workers on Prolific in Studies 1 through 3 (Chapters 2 through 4, participant demographics in Tables 2.2, 3.1, 4.1). Furthermore, these

studies included a questionnaire that asked participants about their experiences and attitudes regarding uptalk, their own voice, and others' voices in public spaces. In this way, analyses in this chapter investigated the ways in which positionality factors interact with one another and influence listeners' recognition of a linguistic feature bearing gendered stereotypical associations.

An examination of these social facets' effects on listeners' linguistic memory supplements existing empirical pursuits of what exactly individuals store in memory from speech input and why (Goldinger 1996; Bradlow & Pisoni 1999). As sociolinguistic work already suggests that components of a listener's demographic identity guide their perception and evaluation of speech input and their speaker (Levon 2014; Walker & Garcia 2014), analyses in this chapter delved deeper into how listener background factors guide how they store and recognize linguistic exemplars.

Across the three memory experiments presented in Chapters 2, 3, and 4, I implemented a survey to examine listeners' positionality regarding the social meanings of uptalk. First, listeners responded to a series of statements regarding their own personal experiences related to their voice. Listeners then indicated their agreement with a series of attitudinal statements about others' voices in public and professional spheres, including the use of uptalk. Listeners' responses were predicted to vary across their distinct positions in their social world: For example, given existing accounts of societal norms that target women's voices with greater linguistic scrutiny (Morgan 2015; Gross 2015; Hachimi 2016), female-identifying listeners were expected to report greater negative experiences with others' perception of their voices in contrast to male-identifying listeners. In turn, these listener background factors were predicted to influence their recognition of rising utterances perceived as uptalk in the old-new judgment task.

The primary goal of this dissertation has been to examine how listeners' ideologically bound expectations of a speaker guide how they recognize that speaker's use of a gender-stereotyped linguistic feature. Studies 1 and 2 examined the extent to which speaker gender, intonational contour, and their interactions with a metalinguistic priming statement in Study 3 were linked to whether listeners recognized utterance productions. Analyses in these studies sought to isolate speaker gender and the use of rising utterances as factors in listeners' old-new judgment performance, regardless of the demographic identities of the listeners. Findings suggested that a speaker's perceived gender did modulate how falling and rising utterances were (falsely) recognized. In this chapter, listeners' own demographic positioning, personal experiences, and attitudes about uptalk were explored and assessed as factors in listeners' old-new recognition. Results and interpretations of listeners' responses in this chapter suggest that the effects of listeners' sociolinguistic background and identity within macrosocial categories on their speech recognition are more complex than previously believed.

5.2. Methods

As discussed in Chapter 2, listeners responded to a sociolinguistic background questionnaire as the final step of each experiment in Studies 1 through 3. Questions were constructed to collect listeners' personal experiences with uptalk, their attitudes about public voices, and their social perceptions of uptalk. The statements in each question were informed by emerging themes found in metalinguistic commentaries as well as among the indexical values previously found embedded in uptalk as a linguistic variable (Tyler 2015; McLemore 1991). Specifically, McLemore (1991) and Warren (2016) found that speakers use uptalk in interlocutor-sensitive contexts to overlay meaning that requests listener confirmation or comprehension as well as

floor-holding or -forfeiting in turn-taking interactions. Responses from Tyler's (2015) study also elucidated other social qualities indexed by a speaker's uptalk use, such as social attractiveness or popularity, as well as insecurity or stupidity. The list below reiterates survey questions described in Chapter 2 (Section 2.3.6) and are henceforth classified as the codified label in brackets:

1. Personal experience questions

Participants responded to questions about their personal sociolinguistic experiences on a scale from 0 to 100 (0 indicating "never" and 100 indicating "all the time") for the frequency of the following experiences:

- [Change] I'm aware of the times I've changed my voice to sound professional or knowledgeable.
- [Sensitive] When I speak, I'm sensitive in making sure my listeners are able to follow what I'm saying.
- [Avoid] I have tried not to use uptalk when speaking.
- [Favorable] I've been told some characteristics of my voice are less favorable.

2. Attitudes about public voices

Listeners then provided their general attitudes about voices in the public sphere by indicating their agreement with the following statements on 7-point Likert scales:

- [Professional] It's important for public image professionals not to use annoying vocal habits.
- [Question] It's annoying when I hear someone sound like they're asking a question even though they're not.
- [Attention] Some qualities about voices make it hard for me to consistently pay attention.

- [Pitch] Voices that are consistently higher pitched are difficult for me to listen to.
- [Uptalk] It's important for public voices like podcast hosts not to use uptalk.

3. Open-response social perceptions of uptalk

Participants provided three unique responses to the prompt: "Uptalk may make the person sound…" to indicate phrases or characteristics they associate with use of uptalk.

5.3. Analyses

5.3.1. Correlations between listener gender, age, and survey responses

The analyses below tested the relationships among each individual question from the sociolinguistic background questionnaire with one another as well as with listener gender and age. Responses to the sociolinguistic background questionnaire along with two demographic elements (age, gender) for all participants across all three studies were first submitted to a series of correlation tests. In order to control for listeners' differential use of the 100-point scale for their responses, participants' responses for all nine questions were transformed into z-scores by participant. For measuring the correlation coefficients for listeners' responses between the nine survey questions (Change, Sensitive, Favorable, Avoid, Professional, Question, Attention, Pitch, Uptalk), listeners' z-scored values, age, and contrast-coded man/woman listener gender categories were submitted to Pearson and Spearman correlation tests. This analysis followed the same correlation tests that Tyler's (2015) study ran to assess the correlations of particular indexical values in listeners' social perceptions of uptalk. By using the same correlation analysis in this chapter, I was able to attend to the strength and direction of the relationship in each of the 110 pairings of all 11 listener factors.

Only participants who self-identified as either men or women were assessed in the coding of participant gender for the correlation matrix due to the statistical limitations of a relatively small sample of non-binary participants recruited across all three studies (N = 5). Participants in Studies 1 through 3 who self-identified as another gender were excluded from correlation analyses. This exclusion was less than ideal due to its criterion based on a binary construction of gender, contradicting previously discussed research that urges sociolinguistic empirical work to explore beyond binary constructions of gender (Zimman 2019; Campbell-Kibler & miles-hercules 2021). In order to investigate effects of participant gender for the vast majority of participants with maximal statistical power, non-binary participants were excluded from the correlation analyses. This yielded an 11x11 correlation matrix on data from the remaining 333 participants, providing the strength and direction of correlations between every listener factor with one another. The matrix was calculated using the rcorr() function in the "corr" R-Studio package (Kuhn, Jackson, & Cimentada 2022). P-values were corrected for multiple interference using Holm's method (Holm 1979). This method prevented potential results with Type 1 errors (e.g. finding a falsely significant correlation) from occurring due to other conflated factors within the relatively large set of comparisons.

5.3.2. Principal components analysis

Listeners' responses to the personal experience and attitude questions were also submitted to a principal components analysis (PCA) in order to test how responses from multiple questions clustered together. Correlation results in Section 5.4 below informed the exclusion of listener gender and age from the PCA, since no significant correlations were found with these demographic factors between any survey question. Responses from all listeners in Studies 1

through 3 were examined in this analysis, including responses from those participants who were previously excluded by their gender identity in the correlations analysis.

PCAs have previously been employed in sociolinguistic analyses that work with multidimensional scaling of social structures and language (e.g. Horvath & Sankoff 1987; Boyd, Hoffman, & Walker 2008; Nagy, Chociej, & Hoffman 2014). For analyses in this dissertation, listeners' responses to the sociolinguistic questionnaire in all three studies were loaded into the PCA to examine which groupings of questions emerged and how the clustering of those questions within principal components best explained variance in responses. The PCA analyzed how much variance in listeners' responses was accounted for by the loading of multiple factors onto a set of principal components. All nine dimensions of listeners' responses were able to be reduced in dimensionality by being applied to a set of principal components in order to observe trends, clusters, and outliers, following Wilson Black et al.'s (2022) application of a PCA in their sociophonetic analysis. In this way, the PCA illuminated the groupings of questions that had the greatest explanatory power to account for listeners' responses along their personal experiences and attitudes about uptalk.

5.4. Predictions

Following broader metalinguistic discourse about uptalk (e.g. Amanatullah & Tinsley 2013), I first predicted that listeners' experiences would vary according to several demographic components in the correlations tests. If listeners' personal attitudes and experiences aligned with popular metalinguistic commentary about uptalk, listeners who identified as women were predicted to differ from those who identified as men. Namely, I predicted that women would indicate that they experienced more self-monitoring and/or policing from others regarding their

use of uptalk than men. This hypothesis would be supported by higher responses for personal experience questions (Change, Sensitive, Avoid, Favorable) from listeners identifying as women versus those identifying as men.

Concerning attitudes about others' use of uptalk, I predicted that younger listeners would exhibit more acceptance of the use of uptalk in public spaces than their older counterparts. Younger speakers and listeners have been shown to lead linguistic, and even more specifically, prosodic change (Milroy & Milroy 1985; Barbieri 2008; Choi, Kim, & Cho 2020). Regarding uptalk, Di Gioacchino & Jessop's (2011) quantitative analysis found that younger speakers were more likely to classify uptalk utterances as more "acceptable" in contrast to older speakers. I thus posited that younger listeners in contrast to older listeners would indicate lower degrees of agreement for attitudinal questions that normalized the social regulation of voices in public spaces (Professional, Question, Pitch, Uptalk).

I also predicted that responses to particular survey questions would correlate with each other. Specifically, I predicted that those questions targeting personal experiences of self-monitoring one's voice (Change, Sensitive, Avoid) would correlate with those asking about experiences of their voice being critiqued by others (Favorable). Those participants who, for example, self-reported little awareness of needing to change their voice in Change may have also exhibited fewer experiences of linguistic policing from others in Favorable.

5.5. Results

5.5.1. Correlations between listener factors

Listeners' responses to specific survey questions were found to correlate with one another, but not with listeners' self-identified gender and age. Positive or negative coefficients with asterisks

in Table 5.1 indicate a significantly positive or negative correlation between two factors, respectively. A positive coefficient with listener gender indicates higher responses from participants identifying as women than from those identifying as men. A negative correlation coefficient with listener gender indicates higher responses from participants identifying as men than those identifying as women. Correlation coefficients among all 11 listener factors are described in Table 5.1 below. In Figure 5.1, correlations in shades of red are in a positive direction; shades of blue denote negative correlations. Brighter colors indicate stronger correlations.

	Gender	Age	Attention	Uptalk	Pitch	Professional	Question	Sensitive	Avoid	Favorable	Change
Change	0.121	-0.118	0.163	0.067	0.048	0.141	0.065	0.420***	0.233**	0.028	1
Favorable	-0.101	-0.004	-0.077	-0.046	-0.029	-0.105	-0.026	-0.232**	0.11	1	
Avoid	0.084	0.059	-0.001	0.224**	-0.052	0.087	0.147	0.331***	1		
Sensitive	0.108	-0.076	0.045	0.073	-0.027	0.165	0.099	1			
Question	0.024	0.172	0.252***	0.547***	0.252***	0.446***	1				
Professional	0.028	0.17	0.188*	0.447***	0.241***	1					
Pitch	0.084	0.174	0.399***	0.264***	1						
Uptalk	0.004	0.121	0.212**	1							
Attention	0.133	0.096	1								
Age	0.072	1									
Gender	1										

*Table 5.1. Correlation coefficients among nine listener variables from the sociolinguistic backgrounds and listener gender and age. N= 333; * = $p < 0.05$, ** = $p < 0.005$, *** = $p < 0.0005$.*

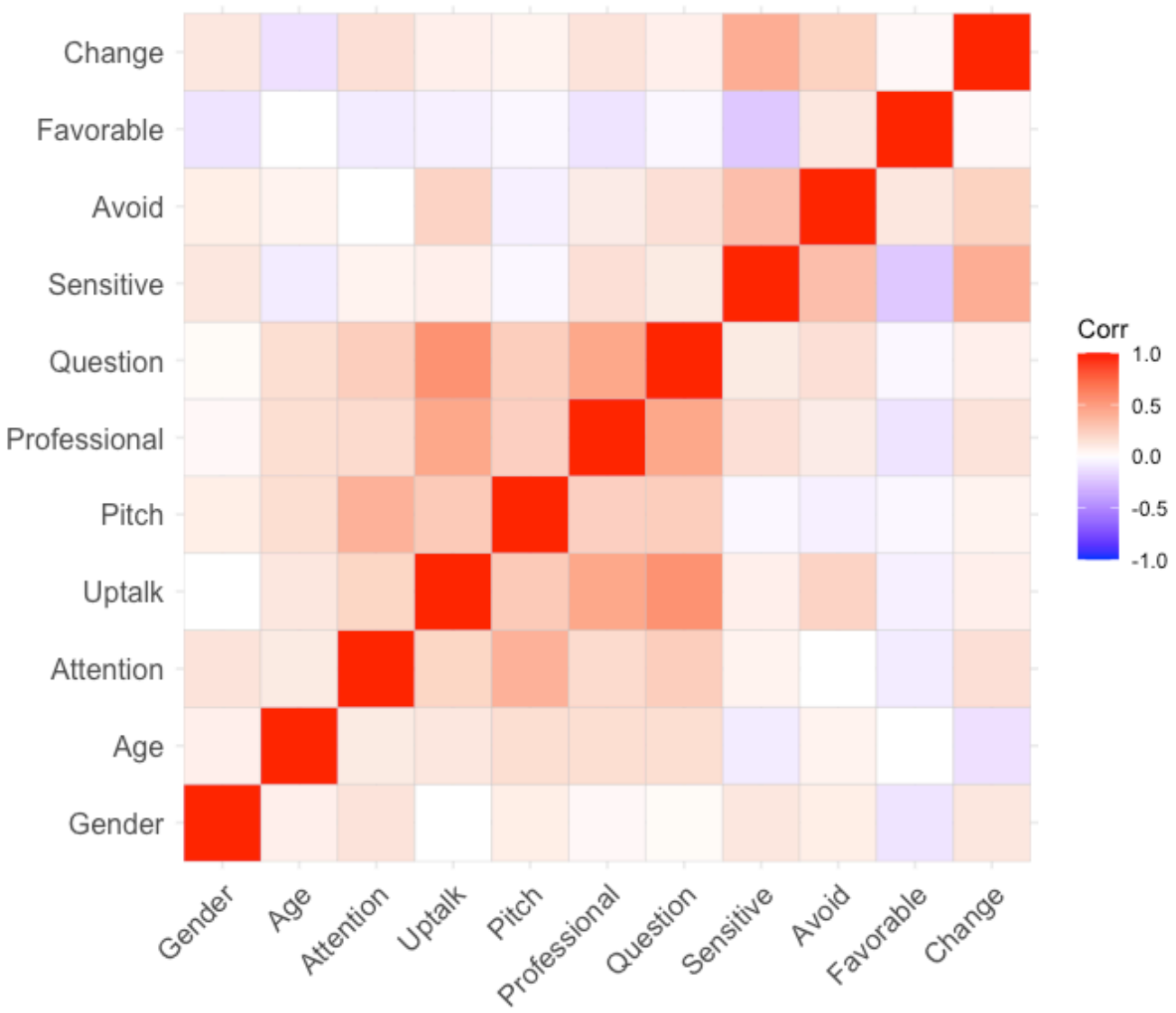

Figure 5.1. Correlation matrix for participant age, gender, and responses to nine sociolinguistic survey questions.

For responses to personal experience questions, I predicted that listeners who identified as women would exhibit higher responses for all questions (Sensitive, Change, Uptalk, Favorable) than listeners who identified as men. However, participants' gender did not significantly correlate with any of these questions.

Listeners' responses to attitude questions were also predicted to be modulated by their age: Specifically, following empirical trends of social evaluations among individuals at different ages (Wollum 2019), agreement with statements indicating negative attitudes toward uptalk was

predicted to increase with their age. However, listener age was not found to significantly correlate with any attitude question or personal experience question (Table 5.1). Listeners' personal experiences and attitudes appeared to fluctuate independently of their gender and age.

Listeners' responses to some of the nine survey questions were found to be significantly correlated with one another. I had previously predicted that listeners who indicated greater degrees of self-monitoring in responses to Change, Avoid, and Sensitive would also indicate greater degrees of linguistic scrutiny in Favorable. Listeners' responses to Sensitive, Change, and Avoid were all significantly correlated with one another. Participants who had relatively high responses to "I'm sensitive in making sure my listeners can follow along" were more likely to report "trying not to use uptalk when speaking" and "changing their voice to sound professional or knowledgeable" in Avoid and Change responses. These trends in participants' responses signified a co-occurrence of self-monitoring and interlocutor sensitivity in personal experiences.

However, Sensitive responses negatively correlated with greater Favorable responses: participants who reported high frequencies of being sensitive to whether their listener could follow along were *less* likely to indicate greater experiences of being told their voice was not "favorable" in contrast to participants with lower Sensitive responses. In other words, being told one's voice is not favorable as a form of linguistic scrutiny did not correspond to greater rates of being sensitive to comprehension of a hypothetical listener. Other survey questions exhibited significant correlations with one another. Listeners' self-monitoring of uptalk, as indicated in their responses to the personal experience question Avoid, significantly correlated with their agreement with the attitudinal statement Uptalk: Listeners who reported greater frequencies of "trying not to use uptalk" were more likely to agree with the statement "It's important for public voices like podcast hosts not to use uptalk." These responses show that listeners who report

greater censoring of their own uptalk use were more likely to hold attitudes against individuals using uptalk in public spaces.

Further, listeners' responses to every attitudinal statement correlated with each other (Professional, Question, Attention, Pitch, Uptalk). If a listener indicated a high degree of agreement with one attitudinal statement, they were more likely to agree with the others, in regards to the following statements: "It's important for public image professionals not to use annoying vocal habits"; "It's annoying when I hear someone sound like they're asking a question even though they're not"; "Some qualities about voices make it hard for me to consistently pay attention"; "Voices that are consistently higher pitched are difficult for me to listen to"; and "It's important for public voices like podcast hosts not to use uptalk."

These results indicate that listeners' attitudes toward other individuals' voices are largely correlated with their personal experiences and introspection about their own voice: Responses of greater frequencies toward personal experience questions correlated with greater agreement toward attitudinal statements for the majority of the sociolinguistic background questions. As the frequency of listeners' reported personal experiences with self-monitoring increased, so did their endorsement of attitudes regulating other voices also facing regulation.

5.5.2. Principal components analysis

To understand how the background questions constituted broader clusters of patterns, as well as to measure how these clusters influenced performance in the old-new recognition task of Study 1, variability of listeners' responses across the nine survey questions as well as their gender and age was reduced to a set of principal components using the *princomp* function in the 'stats' package in R (R Core Team 2012; Harvey & Hanson 2022). The number of principal components was determined by the Scree plot criterion (Braeken & van Assen 2017): The

amount of explained variance was plotted by principal component. Eigenvalues, or the percentages of variance explained, at the 'elbow' of the curve in a Scree plot indicate minimal contribution to the variance in listeners' responses. The Scree plot (Figure 5.2) below illustrates that the first two eigenvalues (principal components 1 and 2) account for 44.5% of the variance in listeners' responses.

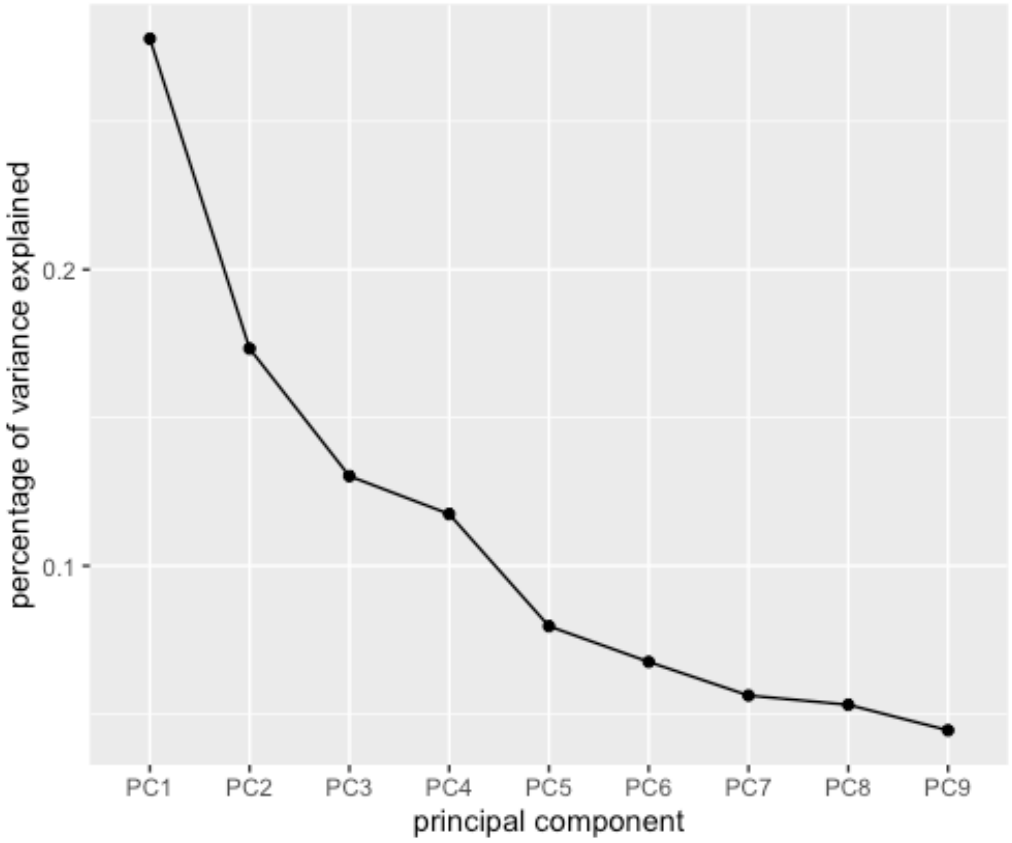

Figure 5.2. Scree plot of ten components' eigenvalues for variances in listeners' sociolinguistic survey responses.

Each survey question was loaded onto two principal components, with the first principal component (PC1) explaining 27.1% of the variance, and the second principal component (PC2) explaining 17.4% of the variance. The loadings of each of the nine variables on the two principal components are included in Table 5.2.

Variable	PC1	PC2
Attention	0.331	-0.204
Uptalk	**0.472**	-0.152
Pitch	0.313	-0.350
Professional	**0.435**	-0.073
Question	**0.464**	-0.155
Sensitive	0.235	**0.584**
Avoid	0.222	**0.452**
Favorable	-0.098	-0.051
Change	0.226	**0.485**

Table 5.2. Loadings for each survey question on principal components 1 and 2.

The two principal components were interpreted by examining questions that had loadings greater than 0.4, according to Stevens' (1992) standard for significance in multivariate analyses. Each loading indicates the relationship between the respective principal component and the original variable (Wilson Black et al. 2022) – in this case, each survey question. The first principal component (PC1) appears to be driven by listeners' responses to three attitude statements: Uptalk ("It's important for public voices like podcast hosts not to use uptalk."), Professional ("It's important for public image professionals not to use annoying vocal habits."), and Question ("It's annoying when I hear someone sound like they're asking a question even though they're not."). The second principal component (PC2) appears to be motivated by personal experience questions Sensitive ("When I speak, I'm sensitive in making sure my listeners are able to follow what I'm saying."), Avoid ("I have tried not to use uptalk when speaking."), and Change ("I'm aware of the times I've changed my voice to sound professional or knowledgeable."). As a result of these characterizations, PC1 can be described as the Expectations factor, and PC2 can be described as the Self-Monitoring factor. The magnitude and directionality of each loading for PC1 and PC2 are illustrated in Figure 5.3, representing each participant's score across both dimensions.

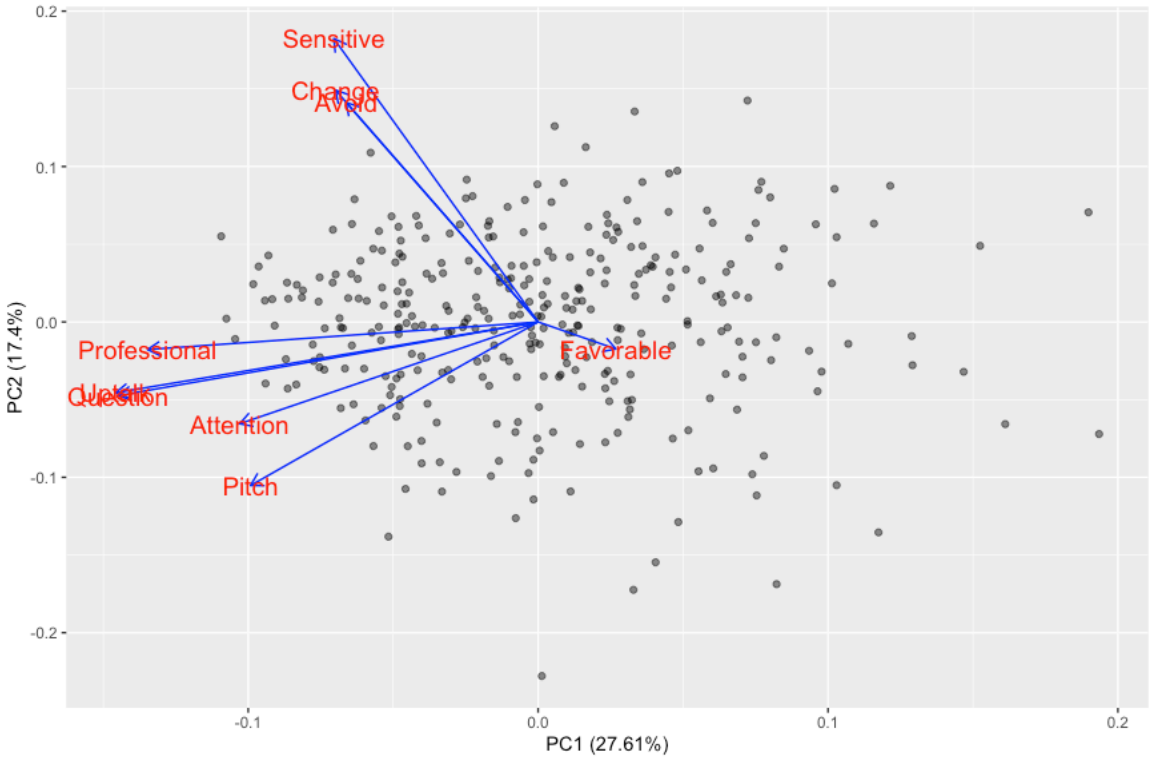

Figure 5.3. Directionality and strength of nine survey questions as loadings in two principal components.

Results from the PCA indicated that responses from attitude and personal experience questions, respectively, clustered together. The loadings for Professional, Uptalk, and Question were similar in strength and directionality, corroborating the significance of the correlations in responses between these questions. The loadings of responses to Sensitive, Change, and Avoid influence PC2. Responses to Favorable, Pitch, and Attention did not exhibit influence on either principal component.

5.6. Interim discussion: listeners' sociolinguistic personal experiences and attitudes

Listeners displayed variation along different dimensions in their personal experiences and attitudes about their own and others' linguistic performances. Regarding their own personal experiences, neither a listener's age nor their gender identity influenced the frequency with which they reported experiences of self-monitoring. Future work that aims to examine the

relationship between an individual's gender and their understanding of uptalk could target other experiences of linguistic subordination, scrutiny, and self-monitoring, such as the frequency with which an individual interrupts or is interrupted, or is criticized for using other gender-linked linguistic features (such as discourse marker 'like' [Schleef 2005; D'Arcy 2007] or creaky voice [Yuasa 2010; Podesva & Szakay 2013]). Previous work in linguistic production (Habasque 2021; Lewandowski 2012) as well as popular metalinguistic commentary (Allen 2019) points to the co-occurrence of features such as 'like' and creaky voice with uptalk, leading to the enregisterment of these features with "Valley Girl Speech" (Ritchart & Arvaniti 2013). The bricolage (Eckert 2018) through which sociolinguistic variables are recruited alongside other socially meaningful forms in linguistic styles may in turn influence listeners' conceptualizations of individual features, such as uptalk, that are embedded in those styles. An examination of other practices along these lines may elucidate influences of gender identity that corroborate metalinguistic trends in which speech styles associated with women's voices face harsher degrees of unsolicited critique and calls for modification (Romm 2017).

Interestingly, the prominence of listeners' personal experiences of linguistic scrutiny and self-monitoring was positively correlated with their endorsement of traditional sociolinguistic attitudes. Listeners with higher responses for personal experiences of scrutiny were more likely to agree with those attitudes that endorse and perpetuate criticism of others' voices. Rather than occurring in an inverted relationship, listeners' stereotypical sociolinguistic attitudes appear to be fortified, rather than weakened, by their own experiences of linguistic self-monitoring and scrutiny. These results suggest that experiences of linguistic scrutiny and self-monitoring act as a stencil for how listeners evaluate others' voices: If an individual exhibits self-monitoring of themselves, to some extent, they expect others to do the same.

This mirroring effect between personal experiences and attitudes toward others corroborates previous social psychological work (Bernheim 1994 inter alia). In particular, individuals have been found to undergo pressures of *social conformity* in models of social exchanges and interactions (Nord 1969). Individuals endeavor to align attitudes and behavior by circulating cultivated social norms (Coleman 2004). Via the *principle of social proof* (Cialdini 1993), individuals are more likely to align with a particular attitude if they have exposure to others doing the same. Listeners' understandings of uptalk may follow a similar route: In their responses to personal experience and attitude questions, listeners' attitudes for standardization of uptalk indicated that they expect others to adhere to the same ideological standards for their own voice. This interpretation invites an empirical lens that captures more experiential and personal factors from listeners and participants rather than larger macro-social categories. Rather than macro-social categories of listeners exhibiting variation in their attitudes and experiences, individuals' own parameters and constraints in their lived experiences regarding self-monitoring and linguistic scrutiny guided the prominence and direction of their attitudes toward others' sociolinguistic practice.

5.7. Listener social factors and old-new recognition

The experiments in Studies 1 through 3 comprised two memory tasks to test participants' 1) speech recognition in the old-new judgment task and 2) content recall in the cloze test. The participant sociolinguistic background and demographic questionnaires were included to contribute to the analyses above as well as to provide one potential option for incorporating nuanced listener factors into the analysis of memory performance. This analysis supported the merging of existing sociolinguistic research that explores individual perceptual differences in

listeners (Loudermilk 2015) with the developing understanding of sociolinguistic perception (Drager & Kirtley 2016). Individuals' own attitudes and experiences have been shown to influence how they interpret others' sociolinguistic practice (e.g. Kinzler & DeJesus 2013), but only a relatively small amount of linguistic research has included these components as meaningful factors in sociolinguistic cognition (e.g. Levon, Buchstaller, & Mearns 2020; Campbell-Kibler 2010). The analysis below examined how listeners' demographic background, personal lived experiences, and attitudes about uptalk influence how they may (mis)remembered rising and falling utterances when listening to speakers perceived as men versus women.

In a post-hoc analysis in the section below, I used responses in the old-new judgment task from Study 1 as an example of an application of particular listener social factors – their age, gender, and existing attitudes and experiences – on their old-new recognition performance. Results from the previous analyses of Study 1 (Chapter 2) showed greater "old" responses for women's rises in contrast to men's rises for both old items (hits) and new items (false alarms). Interpreting these findings, while listeners exhibited a bias to indicate something as "old" rather than as "new," they were more likely to remember rising utterances, especially when produced by a speaker perceived as female versus a speaker perceived as male. This suggests that listeners' memory of uptalk is biased toward recalling women's productions of rises over men's rise productions. To examine these effects more closely, I used the set of Study 1 listeners' old-new responses between speakers perceived as men versus those perceived as women, and I included the listeners' principal component scores from the sociolinguistic background survey as a factor in their performance.

5.7.1. Analysis

To include listeners' existing sociolinguistic background as a predictor of their old-new recognition, listeners' PC1 and PC2 scores, along with their self-reported age and gender were included in mixed effects models fitted to old versus new responses in Study 1. These models also included the factors described in Chapter 2 (old/new status, fall/rise contour, speaker gender, and interaction between contour and speaker gender). Models were then submitted to ANOVAs that assessed model fit with the inclusion of these predictors and their interactions. Specifically, these ANOVA models tested whether the degree of variance in listeners' responses was best accounted for by the inclusion of these factors and their interactions. Speaker, participant, and utterance were included as random intercepts.

5.7.2. Predictions

If listeners' sociolinguistic backgrounds affect the ways in which they (mis)recognize uptalk, then I predicted that their existing attitudes were likely to mediate which falling and rising utterances were remembered from the training task. As described above in the PCA results, PC1 and PC2 were characterized by the clusters of responses from attitudinal statements and personal experience questions, respectively.

Recalling PCA analyses, listeners with greater PC1 scores showed greater agreement with those attitude statements with higher loadings in this principal component, namely Question, Uptalk, and Professional. In other words, listeners with greater PC1 scores were more likely to agree with the following statements: "It's important for public voices like podcast hosts not to use uptalk," "It's important for public image professionals not to use annoying vocal habits," and "It's annoying when I hear someone sound like they're asking a question even though they're not."

Listeners' agreement with these attitudinal statements may be related to their ideological associations they hold between speaker gender and uptalk. Recent metalinguistic commentary illustrates that observations of women's use of uptalk are stereotypically inflated and, at times coincide with prescriptivist standards against the use of this feature in public and professional spheres (Davis 2010; Yoodli 2022). Attitudinal statements in PC1 (Uptalk, Professional, Question) may have in turn received greater agreement from those participants who have upheld stereotypical expectations that women are likely to use uptalk. In this way, listeners with higher PC1 scores (indicating greater endorsement of attitude statements that prescribed standardizing norms for "professional" public voices in the sociolinguistic background survey) may have been more likely to respond "old" for rises when listening to a speaker perceived to be a woman versus a man, as stereotypes related to standardness could be predicted to extend to gendered stereotypes as well. Listeners who exhibited less agreement with these attitudinal statements may not have exhibited this interaction effect to the same extent, since their attitudinal positioning may oppose the standardized regulation of speakers' voices in public spaces according to gender stereotypes. These predicted responses would yield a significant three-way interaction between utterance contour, speaker gender, and listener PC1 score factors.

Returning once more to the PCA results, Listeners with greater PC2 scores were more likely to provide higher responses for the following personal experience questions: Sensitive ("When I speak, I'm sensitive in making sure my listeners are able to follow what I'm saying."), Avoid ("I have tried not to use uptalk when speaking."), and Change ("I'm aware of the times I've changed my voice to sound professional or knowledgeable.").

Listeners' responses to questions loaded onto PC2 may serve as a significant predictor of their old-new judgment performance. In particular, listeners who indicated more frequent

recollections of changing their voice to avoid using uptalk (indicated by a high PC2 score) may have exhibited varying recognition rates for falling and rising utterances compared to those participants with less frequent accounts of these experiences. If these personal experiences of self-monitoring and interlocuter sensitivity led to more sensitivity or attention to others' use of uptalk, then listeners with higher PC2 scores would recognize rising utterances at greater rates than falls, in contrast to those listeners with lower PC2 scores. These predicted results would be supported by a significant interaction between the factors for listeners' PC2 scores and rise/fall utterance contour.

5.7.3. Results

Only those factors that significantly contributed to model fit were included in a final best-fit mixed effects logistic regression predicting listeners' old-new responses. Main effects on old-new recognition responses in Study 1 in the exploratory analysis are included in Table 5.3 below.

Predictor	Estimate	Std. Error	Z-value	P-value
(Intercept)	0.505	0.246	2.050	0.040 *
Old/new status	0.511	0.134	3.802	<0.0001 ***
Fall/rise contour	-0.052	0.405	-0.129	0.897
Speaker gender	0.013	0.113	0.117	0.907
PC1	-0.003	0.056	-0.058	0.954
PC2	0.039	0.072	0.543	0.587
Participant gender	0.839	0.475	1.766	0.077
Participant age	0.005	0.006	0.844	0.399
Status * utterance contour	0.043	0.269	0.159	0.874
Status * speaker gender	-0.252	0.226	-1.116	0.264
Utterance contour * speaker gender	-0.524	0.226	-2.315	0.021 *
Utterance contour * PC1	0.166	0.083	1.988	0.047 *
Utterance contour * PC2	-0.095	0.108	-0.885	0.376
Utterance contour * participant gender	2.314	0.778	2.974	0.003 **
Utterance contour * age	0.014	0.010	1.414	0.157
Participant gender * age	-0.014	0.011	-1.207	0.227
Status * contour * speaker gender	-0.237	0.453	-0.524	0.600
Contour * participant gender * age	-0.040	0.019	-2.158	0.031 *

*Table 5.3. Mixed effects logistic regression summary for Study 1 listeners' old responses. N= 1,536; * = $p < 0.05$, ** = $p < 0.005$, *** = $p < 0.0005$.*

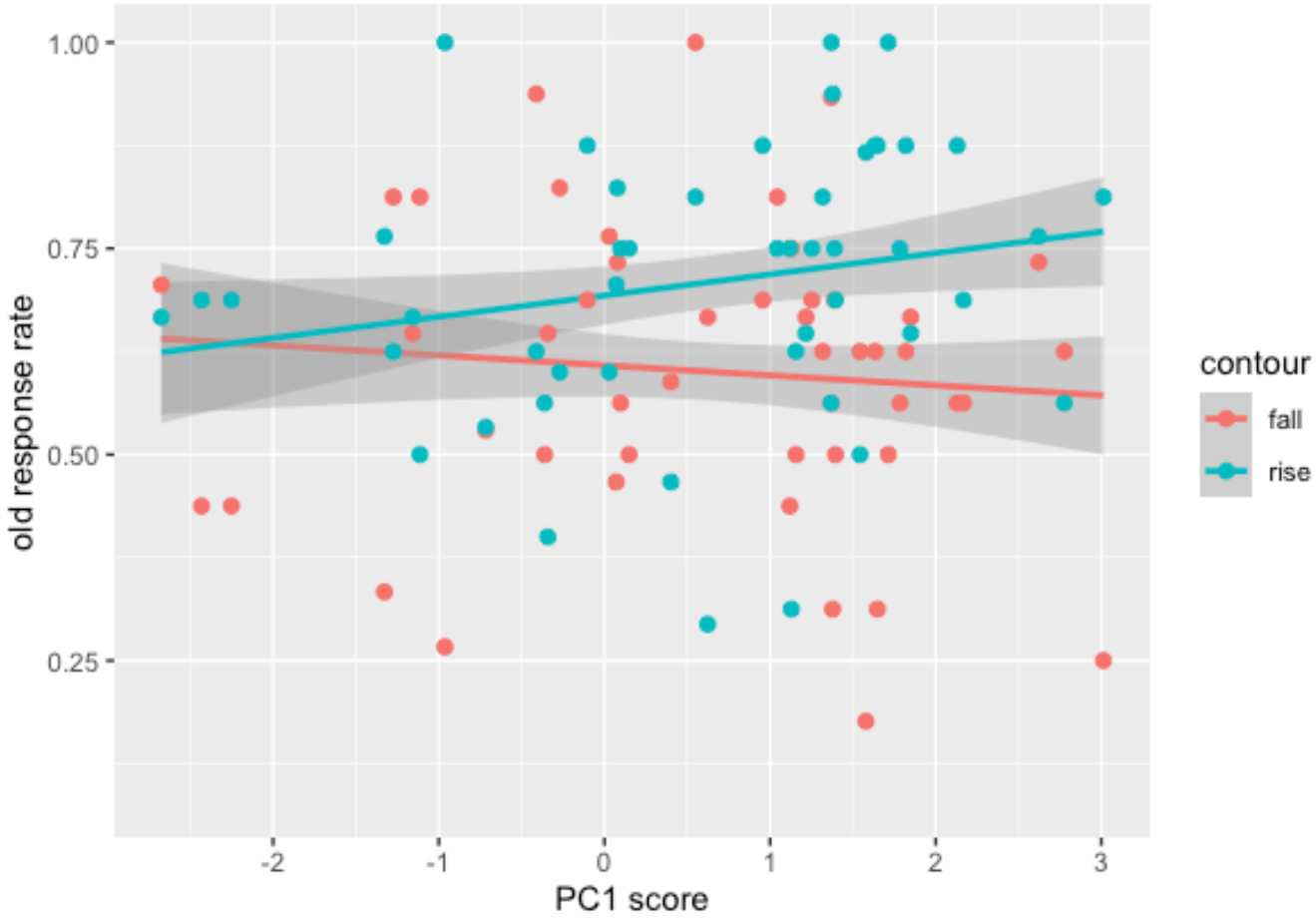

Figure 5.4. Old responses for rises versus falls in Study 1 by participant PC1 score.

Regarding the influence of listeners' background old-new responses, I predicted that listeners bearing attitudes supporting prescriptivist language standardization in the sociolinguistic questionnaire would display a gendered bias in responding "old" to rises more frequently when listening to speakers perceived as women versus speakers perceived as men. Results from the mixed effects logistic regression indicated that the interaction between utterance contour and PC1 scores was a statistically significant predictor of listeners' old-new responses (Table 5.3). Rises were more likely to be classified as "old" by participants with higher PC1 scores, with less of a PC1 score effect with responses for falls. These results suggest that listeners indicating greater agreement with the attitudinal questions with predominant loadings on PC1 (Question,

Professional, Uptalk) were more likely to recognize rises than falls, compared to those participants with lower PC1 scores. Participants who held greater degrees of standardized expectations for the use of uptalk were more likely to classify rising utterances as "old" in contrast to falling utterances. Listeners' responses were not found to be significantly predicted by their PC2 scores nor within any interaction with another factor. Despite my prediction, participants who exhibited greater frequencies of experiencing self-monitoring and interlocutor sensitivity (indicated by a high PC2 score) did not recognize falls or rises at different rates in contrast to low-scoring PC2 participants in the old-new judgment task.

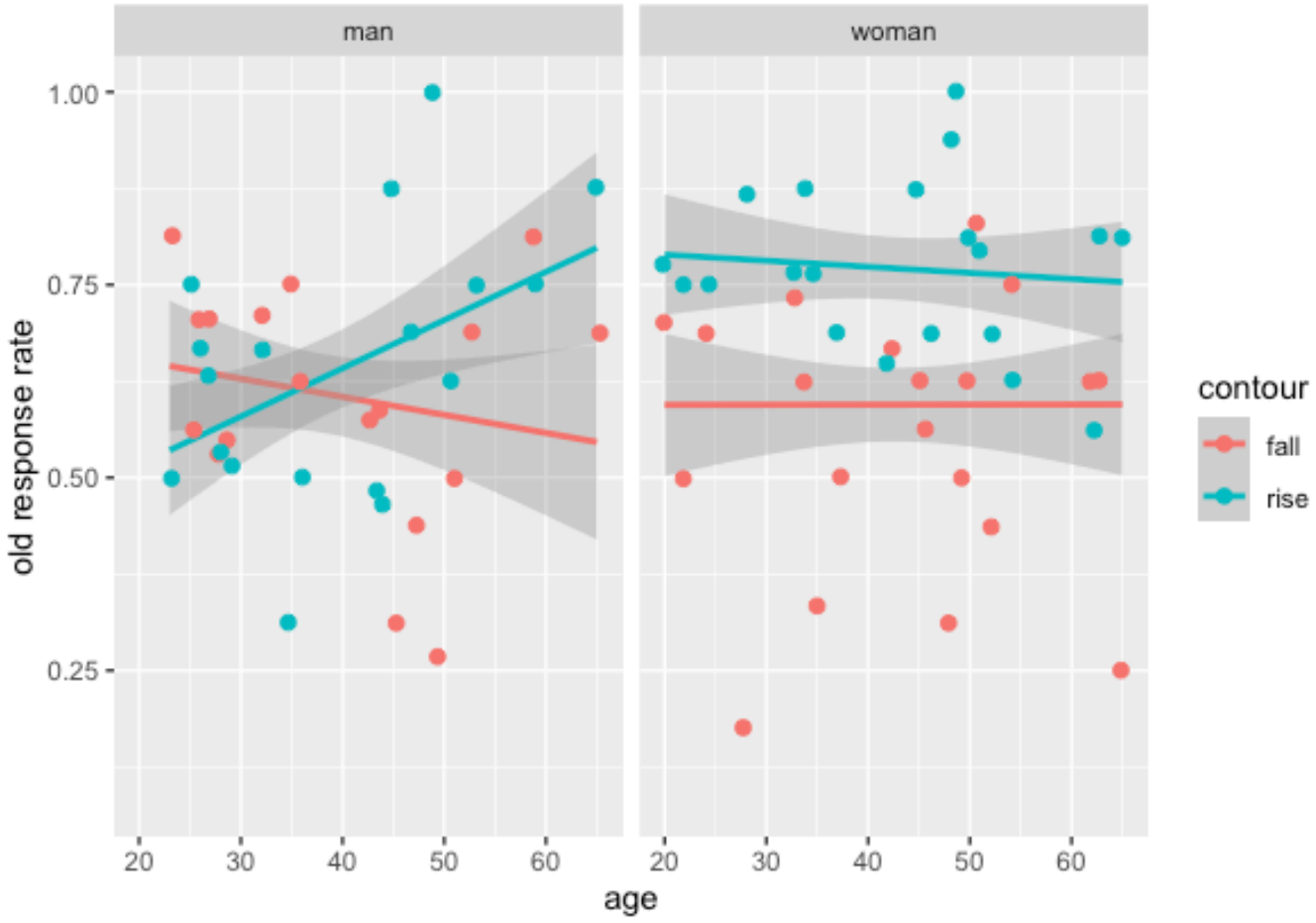

Figure 5.5. Old response rates for rises and falls by participant gender and age.

In addition to the PC1 score effect, listeners' "old" responses also exhibited a significant interaction with utterance contour, listener gender, and listener age. Female participants provided more "old" responses to rising contours over falls than male participants did, indicated by a significantly positive coefficient for the interaction between utterance contour and participant gender (Table 5.3). The influence of participant gender and contour further interacted with participant age: While female listeners of all ages consistently provided more "old" responses for rises than falls, male listeners' age impacted the rate at which they provided old responses for these contours. Rises were most likely classified as "old" by older male listeners in contrast to younger male listeners, with a weaker age effect in the opposite direction for male listeners' memory for falls (Figure 5.4).

These results lend support for my predictions. These effects also occurred independently of the Study 1 main effect of speaker gender. While Study 1 participants had indicated greater "old" responses for rises over falls generally, participants of different genders and ages interacted with this effect differently. The age of a listener only modulated their "old" responses for male listeners. Male listeners' "old" responses to rises increased with their age, while their "old" responses to falls decreased with age. Regardless of their age, female listeners consistently classified rises as "old" more than falls. These results suggest that listeners' performance in old-new speech recognition is not only influenced by the perceived gender of the speaker, but also by listeners' own demographic background—namely, their self-identified gender and age. Additionally, the contour, participant gender, and participant age effects occurred independently of the old/new utterance status effect: Regardless of their own identity, listeners were more likely to provide "old" responses for old items than new items, and they exhibited a bias toward "old" classifications for items of both old and new status. The original significant effects from Study 1

have also remained with the addition of these exploratory factors. Namely, the interaction between speaker gender and utterance contour continued to significantly predict listeners' old responses, in that listeners were more likely to recognize rises from speakers perceived as female versus from those perceived as male.

5.8. Exploratory factors discussion

Components of a listener's sociolinguistic background inform how they produce and interpret socially meaningful language (Eckert & McConnell-Ginet 2003). The ways in which individuals orient themselves around socially meaningful objects, such as sociolinguistic features in practice, consistently shape and re-shape their social positionality. However, in this analysis, listeners' positioning regarding their own personal experiences about changing their voice for others was not found to be predicted by macrosocial patterns of listener age or gender. Listeners' endorsements of attitude statements about standard language and prescriptivism also did not follow these gendered lines, but they did correlate with listeners' personal experiences. Analyses of listeners' sociolinguistic survey responses illustrate that greater rates of personal experiences of monitoring one's voice was met with greater agreement with standard language attitudes by listeners, regardless of their macrosocial identities. In this way, rather than reformulating their expectations for others' voices based on their own experiences of self-monitoring, individuals' attitudes towards others appear to have perpetuated the standards that they and others had for their own voice.

Aspects of listeners' sociolinguistic backgrounds modulated their old-new judgment performance as well. Listeners' PC1 scores, largely comprised by their responses to attitude statements, significantly impacted their "old" responses to utterances of different contours.

Listeners exhibiting greater degrees of agreement with statements about regulating language in public spaces were more likely to classify rises as "old" over falls. Listeners with relatively lower PC1 scores, indicating lower degrees of agreement with the standard language attitude statements, exhibited less of difference in "old" responses between falls and rises. This trend suggests that listeners' attitudes toward uptalk modulate how they remembered speakers' intonational productions in ways that highlight stigmatized features. In this way, the degree to which uptalk bears negative social salience in listeners' attitudes guided how they recognized rising utterances from their speaker. Listeners who held more scrutinizing attitudes toward this feature may have attended to this contour more than listeners with less pertinent negative attitudes about uptalk, which would account for these contours being both better recognized ("old" responses to old items) and falsely remembered ("old" responses to new items). Together, these results illustrate that listeners' positionality toward uptalk mediated how they may have attended to and thus recognized productions of rising utterances.

Independently from their attitudinal positioning, listeners' age and gender also influenced their "old" response rates to falls and rises. Participants identifying as women, in contrast to those identifying as men, provided more "old" responses for rises. Interpreting data from these participants, one's own age and gender identity may inform how they cognitively store episodes of uptalk in larger cognitive representations. Specifically, rising utterances generated more "old" responses from female listeners as well as older male listeners. This intonational feature may bear a greater degree of salience or weight in the perceptual and cognitive linguistic mappings for these listeners. Female listeners, potentially as a consequence of metalinguistic discourse that maps rising utterances to their voices (Ritchart & Arvanati 2014; Quenqua 2021), may have shown greater awareness of uptalk in contrast to those listeners (men) who were not as

frequently the target of stereotypical links with this intonational contour and social groups of speakers.

Responses from male-identifying listeners, contrastively, were mediated by their age. Older male listeners were more likely to classify rises as "old" than falls, with less of a difference in responses between falls and rises for younger male listeners. Trends in "old" responses to rises from older male listeners may indicate that uptalk items are salient intonational productions that garner more attention from these listeners in contrast to younger male listeners, whose social environments or networks may normalize the use of uptalk from others in public spaces (Shokeir 2008). As analyses of metalinguistic prescriptivism (Kickham 2015; Coates 2015), older male individuals' positionings around language change may be oriented more toward linguistic standardization and policing. Trends in metalinguistic commentary on uptalk have been found to be no exception (Davis 2010). In this way, this interpretation of listeners' data considers greater rates of (false) recognition for rises than falls from female and older male listeners. These listener groups may have recognized rising intonation at greater degrees due to their own respective positioning regarding their experiences of self-monitoring (in the case of women) or policing (in the case of older men) of uptalk in their social spheres.

This analysis provided one example of incorporating listener factors into an analysis of sociolinguistic memory. Notions of a listener's social positionality, including existing attitudes, personal experience, and other sociolinguistic background factors (their own gender and age) can continue to be investigated as meaningful factors in the construction and maintenance of sociolinguistic cognitive representations. Further investigation into listener effects in Study 2 and 3 as well as the use of different experimental paradigms would be helpful in more robustly

understanding the ways in which a listener's existing positioning in their sociolinguistic landscape mediates their construction of linguistic cognitive representations.

6. Discussion and conclusion

6.1. Introduction

It has been established that listeners recruit their knowledge of social meaning, as a semiotic property of language, in speech perception (Foulkes & Docherty 2006; Drager 2010; D'Onofrio 2016). As a result, theoretical models accounting for the ways in which listeners encode a linguistic signal into their mental mappings require an inclusion of the listener's ideological positioning and social expectations with regard to their social world. However, the structure of speakers' and listeners' sociolinguistic memory is understudied in relation to overall linguistic cognition (c.f. Vornik, Sharman, & Garry 2003; Knickerbocker & Altarriba 2011; Sumner et al. 2014; Clapp, Vaughn, Todd, & Sumner 2023). Further, predominant variationist approaches (e.g. Labov et al. 2011) explore speakers' socially meaningful linguistic productions and listeners' perception and evaluations of these productions without necessarily engaging the implications for how individuals' linguistic memory relates to this. Of significant exception is recent sociolinguistic work that shows that the encoding and recognition of linguistic forms in memory is facilitated by the interpretation of social information in line with material from the acoustic signal (Sumner & Kataoka 2013). Building on this emerging work that incorporates sociolinguistic meaning into models of cognitive conceptions of language (Foulkes 2010; Sumner et al. 2014; Drager & Kirtley 2016; D'Onofrio 2021), this dissertation has explored listeners' memory as a cognitive window into the ways in which social information, such as the perceived gender identity of a speaker and subsequent gendered expectations, is linked to linguistic mental representations.

To investigate the structures of social expectations in linguistic memory, I examined the sociolinguistic variable uptalk, a feature that has been academically and popularly recognized as

hegemonically associated with women (Tyler 2015; Warren 2016). In three experiments, I tested whether listeners' (false) recognition of this feature exhibited patterning corresponding to the gendered biases prevalent in metalinguistic commentary. First, in Study 1 (Chapter 2), I used an old-new judgment task and cloze test to examine 1) listeners' rates of remembering and falsely remembering intonational rises versus falls and 2) listeners' speech content recall accuracy, both when listening to speakers either perceived as men or women. To evaluate listeners' thresholds of memory for rises more closely, Study 2 (Chapter 3) employed the same experimental paradigms to analyze the extent to which listeners remembered a speaker using rises when this feature was categorically absent in the training passage that listeners had previously heard. I included a priming statement that highlighted metalinguistic links between uptalk and different gendered speaker groups into these paradigms in Study 3 (Chapter 4) in order to inspect the influence of overt ideological reasoning on processes in linguistic recognition. In Chapter 5, I analyzed the ways in which components of a listener's sociolinguistic background constructed different positionalities around uptalk and influenced their performance in Study 1 old-new recognition. In this chapter, I unite findings across the three studies to explore the nature of sociolinguistic cognition. Specifically, I draw on previous postulations regarding listeners' memory of linguistic features, stereotypical links embedded in uptalk, and the influence of individuals' social positioning on their linguistic practice to interpret listeners' sociolinguistic survey responses and their behavior in the memory tasks.

6.2. Linguistic memory and social expectations

6.2.1. Biases in memory load and processing

Between the two memory tasks (old new recognition and cloze test) in Studies 1 through 3, listeners' responses suggest two cognitive preferences. First, when asked to both remember specific intonational tokens as well as specific content words from speech, analyses indicated that their accuracy in content recall was consistently lower than chance levels, unlike their performance in the old-new recognition task. Recalling the initial instructions for the memory tasks, listeners were first told that they would be listening to a podcast passage and would be asked to answer some questions about what they heard in the story. They then were informed that the passage contained sentences spoken with different intonation patterns, and they were subsequently presented with two versions of an utterance produced with contrastive intonational contours. While not within the scope of this dissertation, comparing results between the old-new judgment and cloze tasks suggest a bias to remember intonational tunes of utterances over the specific lexical items that were produced in those utterances. As these instructions in the experimental paradigm aimed to direct listeners' attention to intonational contours in the training phase, this set-up could have caused asymmetrical rates of cognitively encoding intonational productions over content.

Second, listeners' classification of intonational patterns as either "old" or "new" revealed that listeners were more inclined to classify an intonational form in an utterance as one they had heard before as as opposed to one that they had not. This tendency could have been an artifact of the difficulty of the task for the listener. After being asked whether only the intonational contour (rather than the specific content) was the same or different than before, listeners were not statistically more likely to correctly reject new items as 'new' than to classify a new item as

"old" in any of the three studies. Future work enlisting this paradigm could follow other previous studies that operationalize a variable at linguistic levels other than suprasegmental intonational cues (Goldinger 1996; McLennan, Luce, & Charles-Luce 2003; Denby et al. 2018; Sumner & Samuel 2005). As previous work shows variation in listeners' accuracy in speech recognition (Niedzielski 1999; Squires 2013; McGowan 2015; Vaughn & Kendall 2018), listeners' recognition of specific segmental features or specific content items may be better facilitated in this task in contrast to their memory of overall intonational tunes in these utterances. Future work that operationalizes different linguistic features along with those at the prosodic level in memory tasks would further elucidate how different linguistic features are cognitively stored alongside embedded speaker social information.

6.2.2. Memory for rising utterances

As listeners responded to utterances with rising and falling contours in Studies 1 through 3, participants were more likely to classify rises over falls as "old" in Study 1 and Study 2 Block 2. As postulated in Chapter 2, rising contours may have garnered more relative salience in listeners' attention while in the training task, yielding greater rates of "old" responses for both male and female speakers for their rising utterances versus for their falls. However, this interpretation is speculative as these studies measured listeners' recognition rates rather than investigating their attention allotment when first listening to their speaker.

Previous work has shown that rising intonation serves a variety of discourse functions, instantiating social meanings that may be more accessible than when using a falling contour (Warren 2016). As Shokeir (2008) has shown, falling intonation is associated with general confidence, certainty, and completion of one's conversational turn, and has predominantly been classified as the standard intonational structure for declarative utterances in American English

(Rogers 2000). If the "standard" or "idealized" form of declarative utterances is considered to be a falling contour, then the application of rising intonation on this set of utterances is posited to indicate additional functions of that utterance (Cruttenden 1997). Social meaning that emerges from these rising productions in contrast to falls have been shown to indicate a speaker's sensitivity to interlocutor understanding and involvement (Guy & Vonwiller 1984; Allen 1984; Innes 2007), uncertainty (Ward & Hirschberg 1985; Warren 2005), and lack of personal confidence (Gorman 1993).

In Studies 1, 3, and Block 2 of Study 2, listeners' lower rates of recognizing rising utterances over falls suggested that embedded social meaning in these contours guided the encoding, weighting, and recognition of these intonational tokens in listeners' memory. Recalling postulated frequency effects in exemplar structures (Docherty & Foulkes 2014; Drager & Kirtley 2016; Divjak 2019; Gradoville 2023), linguistic forms, including intonational features (Schweitzer 2012) that are heard more frequently lead to stronger cognitive representations that then facilitate encoding and recall of these forms from future linguistic input. Sumner et al. (2014) challenge this notion, citing listeners' relatively greater recognition of tokens of released /t/, a less frequent variant signifying the "canonical" or "idealized" form, in contrast to phonetically reduced variants that are more frequently produced in linguistic production (Ladefoged 1993). Engaging these findings with those of Studies 1, 3, and Study 2 Block 2, listeners were more likely to remember rising contours versus falls, even though falling productions have been found to be more frequently heard in production data in contrast to rises interpreted as uptalk (Fletcher 2005). Listeners in these studies, like those discussed by Sumner et al. (2014), recognized one linguistic variant over another when both were equally presented in an experimental context. As Sumner et al. (2014) posit that the social status that released /t/ has

garnered as a "canonical" variable drove listeners' retrieval of episodes of this token from memory, the social status of rising contours interpreted as uptalk versus falling contours may have manifested similar influences in listeners' recognition of intonational forms. The interaction of the social meanings of rising contours with the perceived gender of the speaker in turn showed influence in listeners' recognition of falling and rising intonation, described below.

6.2.3. Speaker gender effects in (false) recognition of rises

Listeners' responses in the old-new judgment tasks demonstrated that they could, amidst their "old" classification bias, distinguish old intonational items from new, and that social information cued in the speakers' voice informed their performance. In Study 1, in which listeners heard equal rates of rising and falling productions in their training task, and Study 2 Block 1, in which listeners heard only falling utterances and no rises from the speaker, listeners were more likely to classify a rise as "old" – whether or not they actually heard that rise – when they perceived the speaker to be a woman versus as a man. These trends provide further evidence for the social weight of speech exemplars in cognition (Foulkes & Docherty 2010; Sumner et al. 2014). Uptalk as an intonational phenomenon bears salient associations with notions of feminine expressions of gender identity and linguistic styles used by women (Ritchart & Arvanati 2014; Tyler 2015), and metalinguistic discourse on the social expectations and consequences of using this feature underlines the indexical linkage between uptalk and women's voices (Tyler 2015). Listeners' memories of women's rises appeared to be bolstered and weighted by these prevalent social associations. Furthermore, this trend is reminiscent of effects of social biases in linguistic perception: In instantiations of *accent hallucination*, listeners have been shown to falsely recall tokens of non-native English speech from speakers racialized as non-white individuals in contrast to those racialized as white (Babel & Russell 2015; Dovchin 2022). In these cases,

listeners' ideological understandings of non-native English speech conditioned they ways in which they falsely remembered 'accented' speech from different speakers; listeners' false memories of uptalk in the main dissertation studies also appeared to be mediated by ideological expectations of who is likely to have used uptalk in the training phase.

Recognition preferences for female speakers' rises in Study 1 and Study 2 Block 1 also provide support for an influence of binary gender expectations in listeners' mental conceptualizations of uptalk. When presented with groups of speakers who were perceived as either male or female, responses mirrored predominant expectations that ascribe uptalk to women's speech styles (Quenqua 2012; O'Barr & Atkins 1980). However, results from these main studies reflect one portion of the social and perceptual linkage between gender and uptalk, specifically with this feature and those genders most frequently considered in metalinguistic expectations about uptalk. Metalinguistic stereotypes about uptalk pervasively circulate a comparison exclusively between the male-female dichotomy, erasing fluid, less-static notions of gender. As this dissertation tested whether the degree to which listeners' memory of rising utterances followed these prominent stereotypical expectations, my stimuli and analyses were limited to comparing listeners' recognition performance between these binary gender categories. Future perception and memory studies on uptalk could follow Strand's (1999) work by using voices with different degrees of prototypicality for speakers of male and female gender categories. As little experimental work has employed non-binary speakers as stimuli, further study could investigate listeners' recognition of uptalk when the perceived gender of the speaker is not captured by categorical conceptualizations of male versus female speakers. Other potential directions for work delving into the gendered social meanings of uptalk could also build on Tyler's (2015) work by probing into listeners' recognition of uptalk contours when listening to

speakers described as different social types or personae (D'Onofrio 2020) that have imbued gendered and racialized components of meanings in a localized context. These future pursuits could delve into the interaction of nuanced social meanings of uptalk with individuals' mental conceptualizations of this feature.

The second block of Study 2 revealed trends that are opportune for further speculation and investigation. After listeners heard no rises and only falling utterances in Block 1, listeners' responses did not exhibit a significant interaction between speaker gender and utterance contour in Block 2. Rather, rises generated more "old" responses to old items than falls from all speakers, and listeners false alarmed for rises more from men than women (instead of more false alarms for rises from women than men as in Study 1 and Block 1 of Study 2). This trend may have been in relation to its ordering after the training block schema in Block 1. Since the ordering of speaker gender was manipulated across Block 2 participant conditions, participants that heard a woman in Block 1 with no rise productions heard a male speaker produce 50% of his utterances with rising intonation. Rises by men in Block 2 (after no rises in the training passage from a speaker perceived as female) may have been more salient for listeners. Women's rise tokens in Block 2 (after no rises in the Block 1 training passage from a speaker perceived as male) may not have garnered as much surprisal or relative salience. While this justification is speculative and requires further analysis, these results suggest that gendered stereotypes about uptalk can manifest in different ways depending on listeners' previous encounters with uptalk in an experiment. Specifically, half of the participants heard stimuli that violated stereotypical mappings of which gender of speakers is expected to use uptalk by first hearing a female speaker produce no rising intonation, followed by equal amounts of falls and rises produced by a male speaker. The other half of participants heard stimuli that adhered to stereotypical expectations of

uptalk. In this way, the modulation of how many rises a listener heard in one portion of the experiment may have altered the impact of social expectations on their recognition of rise productions in the subsequent block when listeners heard the same rate of rises and falls from speakers. Further investigation could manipulate the ordering of training passages with different frequencies of uptalk or different patternings of this rising tune within the speech stream to explore the cognitive representations of this feature that are active in listeners' memories. These future pursuits could refine existing theories on how listeners' memory of speech is shaped by the frequency with which specific linguistic forms are actually produced by their speaker (Ellis 2002; Abramowicz 2007).

Quantitative accounts of women's relatively greater usage of uptalk (Shokeir 2008; Linneman 2013; Ritchart & Arvaniti 2014) also lend support for the speaker gender effects of listeners' (false) recognition of rises. Previous work has found that listeners' perception of speech and social evaluations of their speaker are affected by the frequency with which a speaker produced socially meaningful or stigmatized linguistic features (Labov et al. 2011; Vaughn & Kendall 2018). This notion entails the listeners' undertaking of *noticing* sociolinguistic features. The act of noticing socially meaningful linguistic forms has been theorized as one component of linguistic perception (Squires 2016), such that noticing, at some degree of awareness, requires listeners' attention to the linguistic signal, and in turn is required for listeners to accurately perceive speech and interpret its imbued social meanings (Drager & Kirtley 2016).

As discussed previously in Chapter 1, individuals' ideologies of language and gender not only guide their attitudes toward gender-stereotyped features, but also modulate the ways in which they produce these features and notice productions of tokens when engaging with other speakers. The stereotype linking women's voices and uptalk may be a product not only of

women's relatively greater usage of this feature (Sando 2009; Ritchart & Arvaniti 2013; Linneman 2013) but also of the social reasoning that interprets this trend by naturalizing uptalk as a component of women's linguistic styles (Tyler 2015). In turn, stereotypes about uptalk may operate like other social stereotypes in the ways that they lead individuals to expect, notice, and evaluate its productions from particular groups of speakers. If participants have noticed greater frequencies of uptalk use from women versus men in their social sphere – potentially due to the prevalence of gender stereotypes with uptalk and/or elevated rates of uptalk usage from women – rising utterances from speakers perceived as women in the main dissertation studies may have contributed to a stronger cognitive representation of uptalk alongside speakers or forms associated with women.

However, listeners in Studies 1 and 3 were presented with equal rates of rises from male and female speakers to control for this frequency effect in the analyses. Listeners, as discussed, bring their own expectations of who is likely to use uptalk, which are formed both by their attitudes and real lived experiences, into their participation in the experiments. Listeners who notice uptalk at greater rates when hearing speakers of a certain gender may have demonstrated stronger activations of these episodic exemplars in cognition, in contrast to those listeners who have different experiences. Listener background analyses in Chapter 5 and metalinguistic responses in Study 3 (Chapter 4) explored further the influence of listeners' attitudes and experiences of using and interpreting uptalk. While the implications for this sociolinguistic link are discussed in sections below, future linguistic work could test listeners' recognition of rises when produced at different frequencies in the training passage for different gendered speakers in order to elucidate this theoretical picture. Further analyses of listeners' recognition of rises when its production is variable in frequency would meaningfully contribute to theoretical postulations

on the effect of token frequency in listeners' cognitive representations of speech (Pierrehumbert 2001).

6.2.4. Metalinguistic priming effects

Listeners in previous studies' perceptual tasks have shown degrees of implicit and explicit awareness of specific components of uptalk's semiotic structure, i.e. its gendered indexical values, as a linguistic feature (e.g. Tyler 2015; Calhoun et al. 2023). The degree to which listeners are overtly cognizant of these gendered meanings of uptalk can vary, and Study 3 (Chapter 4) tested whether I could induce awareness of uptalk as a gendered linguistic feature to modulate listeners' "old" response rates for tokens when hearing rises from speakers of different perceived genders.

Interestingly, the specific gender bias of listeners' metalinguistic prime modulated the effects of other factors between listeners' responses for old and new items. Listeners who received a prime stating that women were most likely to use uptalk were the only ones who more frequently recognized rises over falls, regardless of whether those items were old or new from the training phase. Regarding listeners' responses to old items in particular, the female-biased prime led listeners to exhibit the speaker gender effect from Study 1 and Study 2 Block 1, in which they recognized rises even more frequently than falls when listening to speakers perceived as women versus those perceived as men.

In this way, Study 3 listeners' responses signified an influence of pre-existing ideologies that center uptalk around women in participants' mental representations of this intonational feature. If a prime about uptalk led greater activation strength of cognitive exemplars of rising intonation overall, then primed participants should have exhibited greater (false) recognition rates for rises over falls, regardless of the content of the prime. If a metalinguistic prime that

linked a speaker gender category and uptalk successfully biased listeners to recognize rises from speakers perceived as that targeted gender, then both the male- and female-biased primes should have yielded a significant interaction between speaker gender and utterance contour in their respective directions. The equal prime condition would have potentially yielded equal rates of recognizing rises from speakers perceived as women and as men. Not all of these trends were realized by participants' responses; rather, only the prime that reflected existing stereotypical information about uptalk as linked with women garnered both greater accurate and inaccurate recognition of rises over falls, and yielded more accurate recognition for rises from speakers perceived as the prime-specified gender. These old-new judgments convey stronger activation of uptalk exemplars constructed from listeners' pre-existing linguistic practice and experiences. The explicit priming of uptalk's gendered metalinguistic meaning enhanced listeners' recognition of rising tokens, especially from those speakers perceived as women.

The interaction in recognition between rising utterances and the female-biased prime in Study 3 complicate existing understandings on the power of social meaning at different levels of metalinguistic awareness on listeners' cognitive representations (Levon & Buchstaller 2015; McGowan & Babel 2019). Listeners' differences in old responses between prime conditions in Study 3 convey the influence of metalinguistic themes in listeners' memory that can obscure lived experiences of hearing tokens of rising intonation. In particular, the significant effect of the metalinguistic prime for listeners' responses to new items suggest that overall metalinguistic awareness about women's use of this feature can override listeners' speaker- and utterance-specific memories of uptalk, in that listeners were more likely to falsely recognize rising utterances perceived as uptalk that were previously produced as falls in the training task. In this way, results from Study 3 expose new empirical ground to examine the effects of existing

ideological mappings in listeners' speech recognition, building on emerging work that foregrounds the influence of different degrees of metalinguistic awareness in speech perception (Cheung, Wong, McBride-Chang, Penney, & Ho 2012; Carrera-Sabaté 2014; Harjus 2017; D'Onofrio 2018; McGowan & Babel 2019).

6.3. Listener positionalities and gendered ideology about uptalk

Across all three studies, listeners' interpretations of uptalk as a gendered intonational feature may have encompassed different social meanings. Listeners' experiences regarding uptalk, their own voice, and others, entail nuanced, distinct ideological frames for expectations of who uses this prosodic contour. These existing lenses could have led to variation in the ways that listeners attend to and recall the use of uptalk from different speakers. Analyses in Chapter 5 delved into differences in listeners' social positioning around uptalk. Specifically, I tested whether two components of a listener's positionality – their lived experiences and existing attitudes about others' voices, including evaluations of uptalk – exhibited variation in the listener sample. I probed listeners first for their lived experiences surrounding practices associated with the functions and social meanings of uptalk, specifically their degree of self-monitoring and linguistic scrutiny in interactional contexts. I then asked listeners to indicate their positioning in relation to attitudinal statements that prescribe expectations toward regulating others' voices in their public spheres.

I found that listeners' lived experiences corresponded to their attitudes of standard language ideology. Listeners who reported greater experiences of self-monitoring or having their own voice critiqued were more likely to support standardization of voices and to critique the use of uptalk in public spaces. Macro-social category assignment of listeners by their gender did not

appear to influence listeners' experiences or attitudes in the sociolinguistic background survey. In this way, components of individuals' positionality around uptalk – their understanding of this feature cultivated by their attitudes and the previous experiences of hearing and using this form of rising intonation – appeared to be more strongly correlated with one another than with large-scale identity factors.

These results encourage an empirical lens that recognizes listeners' existing experiences, expectations, and attitudes as interrelated individual components of their sociolinguistic background. Individuals' backgrounds comprise particular experiences of facing pressure (or lack thereof) to present an "appropriate" public voice, and their reactions in the ways they monitor their own voice and construct attitudes about others. Awareness of and attitudes toward the self and others may in turn affect how incoming and future socially meaningful language is interpreted.

As a jumping off point for further work, I used the sample of Study 1 listeners to test the effect of listeners' background on their old-new recognition. Analyses found that three facets of a listener's identity affected how they remembered rising utterances: their self-reported age, gender, and standard language attitudes (reflected by their PC1 score). Listeners identifying as women were more likely to recognize rises over falls from speakers; listeners identifying as men exhibited variation by age. Older male listeners recognized rises more than falls, with the opposite trend (more recognition for falls than rises) for younger male listeners. Some of these findings can be considered in line with existing variation in production and evaluations of uptalk from these gendered groups of listeners. Regarding age effects in old-new recognition, further analyses are required to elucidate the influence of listener age and evaluations of speech. While Di Gioacchino & Jessop's (2010) work found that older listeners found uptalk tokens more

'jarring' and 'unacceptable' than younger listeners this trend was only investigated in listeners' speech evaluations, not including recognition of linguistic forms in cognition.

An interpretation of listener gender effects on their old-new recognition requires a review of previous work on the relationship between gender identity and evaluations of uptalk. Specifically, Shokeir (2008) found that South Ontario Canadian male listeners heard falling utterances as more "finished," "certain," and "confident" than utterances with terminal rises, and female listeners did not distinguish these different categories of contours along these dimensions of meaning. Furthermore, Levon's (2016) study of speakers' production of rising declaratives in London illustrated differences in functions for which men and women recruited uptalk into their styles. Whereas male speakers' primary motivations for uptalk use derived from their desire to remain in conversational focus with a mixed-gender interlocutor group, female speakers' productions emerged from their goal to avoid conflict or interpersonal disagreement. These previous analyses provide support for gender-linked variation in listeners' understandings of the functions that uptalk serves, and their interpretations of its discursive meanings. Listeners identifying as men or women varied in their recognition rates of rising and falling utterances potentially due to differences in the ways they interpret and map rising and falling contours to different dimensions of social meaning.

Listeners who expressed greater alignment with their attitudes and standard language ideology exhibited greater recognition for rises than falls, in contrast to those listeners with less standard language-linked attitudes. This trend lends support for the influence of attitudinal positioning on listeners' recognition of speech. Listeners bearing attitudes that attend to prescribing how speakers "ought" to sound in public spaces may have recognized rising utterances more than falls due to this attitudinal stance that marks uptalk as a deviant form of

speech. In this way, listeners' attitudes against uptalk in social spaces did not lead them to ignore instances of rising intonation from their speakers; rather, this attitude appeared to facilitate their recognition of this feature. This interpretation can be supported by previous work in sociolinguistic perception: Using the sociolinguistic variable (ING) as an example, the social stigmatization of the alveolar variant *–in* versus its velar variant *–ing* has cyclically garnered more metalinguistic discourse and greater degrees of awareness of this feature. As Levon & Fox (2014) posit, without the necessary degree of social awareness of this variant's social status, listeners' social evaluations may not be affected by the frequency with which a speaker produces alveolar tokens over velar tokens. In regards to listeners' attitudes and recognition of uptalk in the exploratory analyses, the prevalence of uptalk's negatively aligned social meanings in listeners' attitudes appeared to affect the rate at which they (falsely) recognized rising and falling productions in the old-new judgment task. In this way, listeners' endorsement of negative statements about uptalk in their attitude responses may have led to greater rates of recognition for rises interpreted as uptalk in contrast to falling utterances.

These empirical findings from the exploratory analysis provide more evidence that a listener's gender, age, and the social positionings entailed by these components of identity, guide their interpretations of uptalk. The variation in listeners' memory of rises correspond to differences in how they encode and conceptualize uptalk in their mental mappings of their sociolinguistic landscape. Components of a listener's demographic identity and attitudes are some dimensions along which they are positioned and position themselves in relation to others in their interactions and interpretations of sociolinguistic practice, and these experiences guide the ways in which sociolinguistic input is cognitively represented. In this way, the analyses in this dissertation sought to include components of a listener's sociolinguistic reasoning as a relevant

frame for their encoding and conception of linguistic cognitive representations that are at work in these interactional processes. Listeners' performance in recognizing old and new intonational contours reflect that stereotypical expectations about uptalk can mediate (false) memories of this feature, but that components of social identity, such as gender, age, and language standardization attitudes, modulate the ways in which listeners recognize this gendered intonational contour over falling contours.

6.4. Conclusion

The primary goal of this dissertation was to provide a more thorough understanding and investigation of listeners' sociolinguistic memory. While individuals exhibit rigorous command and application of their social reasoning in linguistic production and processes in perception (Strand 1999; Eckert & McConnell-Ginet 2003; Cambell-Kibler 2006), less academic work has directly engaged performance of listeners' memory which crucially hold sources of this sociolinguistic knowledge. In an examination of listeners' memory of old and new intonational information, components of a listener's expectations about a speaker governed how a linguistic feature, uptalk, was encoded in connection with other social meanings. Listeners' recognition performance was biased toward rises produced by speakers perceived as female versus those as male, suggesting that social expectations of male and female speakers facilitated the recognition of intonational contours from memory. Metalinguistic primes that mirrored predominant social expectations about uptalk prompted listeners to recognize rising utterances from all speakers, illustrating the influence of stereotypical information on listeners' attention and recognition of rising contours. The interpretations of these trends built on emerging sociolinguistic research that theorizes the ways in which social meaning fortifies the activation of linguistic concepts in the

mind (Sumner et al. 2014; McGowan 2015; Babel & Russell 2015; D'Onofrio 2015). While this cognitive meaning-making work is often automatic in perception and other sites of linguistic praxis (Schilling-Estes 1998; Labov et al. 2011), they are all the more vital in the very perceptible moves that listeners and speakers make with these mental capacities. As speakers and listeners are both accountable in shaping language across time and space, they crucially use and shift their views of their world in order to do so. The interpretation and meaning emerging from language derive from the space between a speaker's linguistic expressions and a listener's interpretation of this signal, with both actors' knowledge of sociolinguistic forms as a guide.

References

Abramowicz, Ł. (2007). Sociolinguistics meets exemplar theory: Frequency and recency effects in (ing). *University of Pennsylvania Working Papers in Linguistics*, *13*(2), 3.

Agha, A. (2003). The social life of cultural value. *Language & Communication*, *23*(3-4), 231-273.

Ainsworth, J. E. (1993). In a Different Register: The Pragmatics of Powerlessness in Police Interrogation. *The Yale Law Journal*, 103(2), 259-322.

Albarracin, D., & Shavitt, S. (2018). Attitudes and attitude change. *Annual Review of Psychology*, *69*, 299-327.

Alderton, R. (2020). Speaker gender and salience in sociolinguistic speech perception: GOOSE-fronting in Standard Southern British English. *Journal of English Linguistics*, *48*(1), 72-96.

Allen, K. (1984). The component functions of the HRT contour in Australian declarative sentences. *Australian Journal of Linguistics*, 4: 19-32.

Allen, S. (2019). Uptalk? and Vocal Fry? An Epidemic? *Grammarly Blog*. https://www.grammarly.com/blog/uptalk-and-vocal-fry-an-epidemic/

Amanatullah, E. T., & Tinsley, C. H. (2013). Ask and ye shall receive? How gender and status moderate negotiation success. *Negotiation and Conflict Management Research*, *6*(4), 253-272.

Anderson, R. C., Klofstad, C. A., Mayew, W. J., & Vekatachalam, M. (2014). Vocal fry may undermine the success of young women in the labour market. *PLoS One*, 9(5), e97506.

Angouri, J. (2021). Introduction: Language, gender, and sexuality: Sketching out the field. In *The Routledge Handbook of Language, Gender, and Sexuality* (pp. 1-22). Routledge.

Asch, M., & Brogan, F. D. (2022). Uptalk in Chicano Southern California English. *American Speech: A Quarterly of Linguistic Usage*, 1-27.

Aslan, E., & Vásquez, C. (2018). 'Cash me ousside': A citizen sociolinguistic analysis of online metalinguistic commentary. *Journal of Sociolinguistics*, *22*(4), 406-431.

Augoustinos, M., & Walker, I. (1998). The construction of stereotypes within social psychology: From social cognition to ideology. *Theory & Psychology*, *8*(5), 629-652.

Austen, M., & Campbell-Kibler, K. (2022). Real-time speaker evaluation: How useful is it, and what does it measure? *Language*, *98*(2), e108-e130.

Babel, A. M. (2016). Silence as control: Shame and self-consciousness in sociolinguistic positioning. *Awareness and Control in Sociolinguistic Research*, 200-227.

Babel, M. (2010). Dialect divergence and convergence in New Zealand English. *Language in Society*, *39*(4), 437-456.

Babel, M. E. (2009). *Phonetic and social selectivity in speech accommodation*. University of California, Berkeley.

Babel, M., & Russell, J. (2015). Expectations and speech intelligibility. *The Journal of the Acoustical Society of America*, *137*(5), 2823-2833.

Baddeley, A. (1992). Working memory: The interface between memory and cognition. *Journal of Cognitive Neuroscience*, *4*(3), 281-288.

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*(3), 189-208.

Baldoni, J. (2015, July 30). Will "Upspeak" Hurt Your Career? *Forbes*. https://www.forbes.com/sites/johnbaldoni/2015/07/30/will-upspeak-hurt-your-career/

Barbieri, F. (2008). Patterns of age-based linguistic variation in American English. *Journal of Sociolinguistics*, *12*(1), 58-88.

Barreda, S. (2017). An investigation of the systematic use of spectral information in the determination of apparent-talker height. *The Journal of the Acoustical Society of America*, *141*(6), 4781-4792.

Baxter, J. (2003). *Positioning gender in discourse: A feminist methodology*. Springer.

Beckman, M. E., Hirschberg, J., & Shattuck-Hufnagel, S. (2006). The original ToBI system and the evolution of the ToBI framework. *Prosodic Typology: The Phonology of Intonation and Phrasing*, *1*(9).

Bernheim, B. D. (1994). A theory of conformity. *Journal of Political Economy*, *102*(5), 841-877.

Bialecki, J. (2021). Coda: An Origin Story (An essay in the Self-Positionality in the Anthropology of Christianity series). *New Directions in the Anthropology of Christianity*, *17*.

Bing, J., & Bergvall, V. (1996). (1996)" The Question of Questions: Beyond Binary Thinking." In *Rethinking Language and Gender Research.* London: Longman, 1-30.

Bogetić, K. (2016). Metalinguistic comments in teenage personal blogs: Bringing youth voices to studies of youth, language and technology. *Text & Talk*, *36*(3), 245-268.

Bourdieu, P. (1977). ‘‘Cultural Reproduction and Social Reproduction.’’ In *Power and Ideology in Education*, J. Karabel and A Halsey (Eds). New York: Oxford University Press.

Bourhis, R. Y., & Giles, H. (1977). The language of intergroup distinctiveness. *Language, Ethnicity and Intergroup Relations*, *13*, 119.

Boyd, S., Hoffman, M. F., & Walker, J. A. (2015). Sociolinguistic variation among multilingual youth: comparing Swedish cities and Toronto. *Language, Youth and Identity in the 21st Century: Linguistic Practices Across Urban Spaces*, 290e305.

Bradlow, A. R., & Pisoni, D. B. (1999). Recognition of spoken words by native and non-native listeners: Talker-, listener-, and item-related factors. *The Journal of the Acoustical Society of America*, *106*(4), 2074-2085.

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. I. (1999). Training Japanese listeners to identify English/r/and/l: Long-term retention of learning in perception and production. *Perception & Psychophysics*, *61*, 977-985.

Bradlow, A. R., Nygaard, L. C., & Pisoni, D. B. (1999). Effects of talker, rate, and amplitude variation on recognition memory for spoken words. *Perception & Psychophysics*, *61*(2), 206-219.

Braeken, J., & Van Assen, M. A. (2017). An empirical Kaiser criterion. *Psychological Methods*, *22*(3), 450.

Bridges, J. (2017). Gendering metapragmatics in online discourse: "Mansplaining man gonna mansplain…". *Discourse, Context & Media*, *20*, 94-102.

Bucholtz, M. (2009). From stance to style: Gender, interaction, and indexicality in Mexican immigrant youth slang. *UC Santa Barbara Previously Published Works.*

Bucholtz, M. (2019). Sociolinguists trying to make a difference: Race, research, and linguistic activism. In *The Routledge Companion to the Work of John R. Rickford* (pp. 242-250). Routledge.

Bucholtz, M. (2019). The public life of white affects. *Journal of Sociolinguistics*, *23*(5), 485-504.

Bunting, M., Cowan, N., & Colflesh, G. (2008). The deployment of attention in short-term memory tasks: Trade-offs between immediate and delayed deployment. *Memory & Cognition*, *36*(4), 799-812.

Butler, J. (1990)/ Gender trouble, feminist theory, and psychoanalytic discourse. *Feminism/Postmoderdism 327*, 324-340.

Butler, J. (2013). Burning acts: Injurious speech. In *Performativity and Performance* (pp. 197-227). Routledge.

Calder, J. (2019). The fierceness of fronted/s: Linguistic rhematization through visual transformation. *Language in Society*, *48*(1), 31-64.

Calder, J. (2020). Language, gender and sexuality in 2019: interrogating normativities in the field. *Gender & Language*, *14*(4).

Calder, J., & Steele, A. (2019). Gender in sociolinguistic variation beyond the binary. Paper presented at Linguistic Society of America Annual Meeting, New York, NY.

Calhoun, S., Warren, P., Mills, J., & Agnew, J. Uptalk and the frequency code: how gender affects iconic associations of pitch. Paper presented at International Congress of Phonetic Sciences, Prague, Czech Republic.

Cameron, D. (2015). Evolution, Language and the Battle of the Sexes: A Feminist Linguist Encounters Evolutionary Psychology. *Australian Feminist Studies*, *30*(86), 351-358.

Cameron, D., & Kulick, D. (2003). *Language and sexuality*. Cambridge University Press.

Campbell-Kibler , K. and miles-hercules , D. 2021). Perception of gender and sexuality. In: *The Routledge Handbook of Language, Gender, and Sexuality* (eds. J. Angouri and J. Baxter), 52-68. London: Routledge .

Campbell-Kibler, K. (2006). Variation and the listener: The contextual meanings of (ING). *University of Pennsylvania Working Papers in Linguistics*, *12*(2), 6.

Campbell-Kibler, K. (2007). Accent, (ING), and the social logic of listener perceptions. *American Speech*, *82*(1), 32-64.

Campbell-Kibler, K. (2008). I'll be the judge of that: Diversity in social perceptions of (ING). *Language in Society*, *37*(5), 637-659.

Campbell-Kibler, K. (2009). The nature of sociolinguistic perception. *Language Variation and Change*, *21*(1), 135-156.

Campbell-Kibler, K. (2010). The effect of speaker information on attitudes toward (ING). *Journal of Language and Social Psychology*, *29*(2), 214-223.

Campbell-Kibler, K. (2010). Sociolinguistics and perception. *Language and Linguistics Compass*, *4*(6), 377-389.

Campbell-Kibler, K. (2016). Towards a cognitively realistic model of meaningful sociolinguistic variation. In *Awareness and Control in Sociolinguistic Research*, 123-151.

Carmichael, K. (2016). Place-linked expectations and listener awareness of regional accents. In *Awareness and Control in Sociolinguistic Research*, 152-176.

Carrera-Sabaté, J. (2014). Does meta-linguistic awareness play any role at the beginning of an ongoing sound change? The case of some vowel-ended verbs in Catalan. *Sociolinguistic Studies*, *8*(2), 193.

Casasanto, L. S. (2009). How Do Listeners Represent Sociolinguistic Knowledge?. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 31, No. 31).

Chao, M., & Bursten, J. (2016). Heard but Misunderstood: Understanding Negative Reactions to Female Vocal Fry.

Chao, M., & Bursten, J. R. (2021). Girl talk: Understanding negative reactions to female vocal fry. *Hypatia*, *36*(1), 42-59.

Cheung, H., Chung, K. K. H., Wong, S. W. L., McBride-Chang, C., Penney, T. B., & Ho, C. S.-H. (2010). Speech perception, metalinguistic awareness, reading, and vocabulary in Chinese-English bilingual children. *Journal of Educational Psychology*, 102, 367–380.

Chevrot, J. P., Drager, K., & Foulkes, P. (2018). Editors' introduction and review: Sociolinguistic variation and cognitive science. *Topics in Cognitive Science*, *10*(4), 679-695.

Choi, J., Kim, S., & Cho, T. (2020). An apparent-time study of an ongoing sound change in Seoul Korean: A prosodic account. *PloS One*, *15*(10), e0240682.

Chun, M. M., & Turk-Browne, N. B. (2007). Interactions between attention and memory. *Current Opinion in Neurobiology*, *17*(2), 177-184.

Church, B. A., & Schacter, D. L. (1994). Perceptual specificity of auditory priming: implicit memory for voice intonation and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(3), 521.

Clapp, W., Vaughn, C., Todd, S., & Sumner, M. (2023). Talker-specificity and token-specificity in recognition memory. *Cognition*, *237*, 105450.

Clifford, L. (2019). Late acquisition of gendered phonetics: voice feminization in transgender women. Paper presented at New Ways of Analyzing Variation 48, Eugene, OR.

Coates, J. (2015). *Women, men and language: A sociolinguistic account of gender differences in language*. Routledge.

Cole, J., & Shattuck-Hufnagel, S. (2011). The phonology and phonetics of perceived prosody: What do listeners imitate?. In *Twelfth Annual Conference of the International Speech Communication Association*.

Coleman, S. (2004). The effect of social conformity on collective voting behavior. *Political analysis*, *12*(1), 76-96.

Couper-Kuhlen, E., & Selting, M. (1996). Towards an interactional perspective on prosody and a prosodic perspective on. *Prosody in Conversation: Interactional Studies*, *11*.

Craik, F. I., & Kirsner, K. (1974). The effect of speaker's voice on word recognition. *The Quarterly Journal of Experimental Psychology*, *26*(2), 274-284.

Cruttenden, A. (1981). Falls and rises: meanings and universals. *Journal of Linguistics*, *17*(1), 77-91.

Cruttenden, A. (1997). *Intonation*. Cambridge University Press.

Cutler, A., Dahan, D., & Van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, *40*(2), 141-201.

D'Arcy, A. (2007). Like and language ideology: Disentangling fact from fiction. *American Speech*, *82*(4), 386-419.

D'Imperio, M. (2000). *The role of perception in defining tonal targets and their alignment*. PhD dissertation, The Ohio State University.

D'Onofrio, A. (2015). Persona-based information shapes linguistic perception: Valley Girls and California vowels. *Journal of Sociolinguistics*, *19*(2), 241-256.

D'Onofrio, A. (2016). Social meaning in linguistic perception. PhD Dissertation, Stanford University.

D'Onofrio, A. (2018). Controlled and automatic perceptions of a sociolinguistic marker. *Language Variation and Change*, *30*(2), 261-285.

D'Onofrio, A. (2018). Personae and phonetic detail in sociolinguistic signs. *Language in Society*, *47*(4), 513-539.

D'Onofrio, A. (2019). Complicating categories: Personae mediate racialized expectations of non-native speech. *Journal of Sociolinguistics*, *23*(4), 346-366.

D'Onofrio, A., & Benheim, J. (2020). Contextualizing reversal: Local dynamics of the Northern Cities Shift in a Chicago community. *Journal of Sociolinguistics*, *24*(4), 469-491.

D'Onofrio, A., Hall-Lew, L., Moore, E., & Podesva, R. J. (2021). Sociolinguistic signs as cognitive representations. In *Social meaning in linguistic variation: Theorizing the third wave.* Cambridge University Press.

Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology*, *44*(1), 20.

Davis, H. (2020, January 29). The Uptalk Epidemic. *Psychology Today*. https://www.psychologytoday.com/us/blog/caveman-logic/201010/the-uptalk-epidemic

Davis, J. L., & Hall, K. (2021). Ethnography and the Shifting Semiotics of Gender and Sexuality: Practice, Ideology, Theory. In *The Routledge Handbook of Language, Gender and Sexuality* (pp. 93-107). Routledge.

De Houwer, J., Thomas, S., & Baeyens, F. (2001). Association learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin*, *127*(6), 853.

Denby, T., Schecter, J., Arn, S., Dimov, S., & Goldrick, M. (2018). Contextual variability and exemplar strength in phonotactic learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(2), 280.

Di Gioacchino, M., & Jessop, L. C. (2010). Uptalk-towards a quantitative analysis. *Toronto Working Papers in Linguistics*, *33*.

Divjak, D. (2019). *Frequency in language: Memory, attention and learning*. Cambridge University Press.

Dodsworth, R. (2008). Sociological consciousness as a component of linguistic variation. *Journal of Sociolinguistics*, *12*(1), 34-57.

Dovchin, S. (2022). Translingual discrimination. In *Elements in Intercultural Communication*. Cambridge University Press.

Drager, K. (2005). From bad to bed: the relationship between perceived age and vowel perception in New Zealand English. *Te Reo 48*, 55-68 .

Drager, K. (2010). Sociophonetic variation in speech perception. *Language and Linguistics Compass 4*(7):473-480.

Drager, K. (2011). Speaker age and vowel perception. *Language and Speech 54*(1):99-121.

Drager, K. (2016). Constructing style: phonetic variation in discursive functions of like. In Heike Pichler (Ed.) *Discourse-Pragmatic Variation and Change in English: New Methods and Insights*, Cambridge University Press, 232-251.

Drager, K., & Kirtley, M. J. (2016). Awareness, salience, and stereotypes in exemplar-based models of speech production and perception. *Awareness and Control in Sociolinguistic Research*, 1-24.

Drager, K., Hardeman-Guthrie, K., Schutz, R., & Chik, I. (2021). Perceptions of Style: A Focus on Fundamental Frequency and Perceived Social Characteristics. In *Social Meaning and Linguistic Variation: Theorizing the Third Wave*, 176.

Dziura, J. (2020). Hey Ladies—Here's How to Ignore All the Critics and Use Upspeak to Your Advantage. *The Muse*.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace Jovanovich College Publishers.

Eagly, A. H., Chen, S., Chaiken, S., & Shaw-Barnes, K. (1999). The impact of attitudes on memory: An affair to remember. *Psychological Bulletin*, *125*(1), 64.

Eagly, A. H., Kulesa, P., Chen, S., & Chaiken, S. (2001). Do attitudes affect memory? Tests of the congeniality hypothesis. *Current Directions in Psychological Science*, *10*(1), 5-9.

Eckert, P. (1989a). *Jocks and burnouts: Social categories and identity in the high school*. Teachers College Press.

Eckert, P. (1989b). The whole woman: Sex and gender differences in variation. *Language Variation and Change*, *1*(3), 245-267.

Eckert, P. (2003). Language and adolescent peer groups. *Journal of Language and Social Psychology*, *22*(1), 112-118.

Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics*, *12*(4), 453-476.

Eckert, P. (2011). Where does the social stop. *Language Variation-European Perspectives III*, 13-29.

Eckert, P. (2012). Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology*, *41*, 87-100.

Eckert, P. (2014). Language and gender in adolescence. *The Handbook of Language, Gender, and Sexuality*, 529-545.

Eckert, P. (2017). Age as a sociolinguistic variable. *The Handbook of Sociolinguistics*, 151-167.

Eckert, P., & McConnell-Ginet, S. (1992). Think practically and look locally: Language and gender as community-based practice. *Annual Review of Anthropology*, *21*(1), 461-488.

Eckert, P., & McConnell-Ginet, S. (2003). *Language and Gender*. Cambridge University Press.

Eckert, P., & Podesva, R. J. (2021). Non-binary approaches to gender and sexuality. *The Routledge Handbook of Language, Gender, and Sexuality*, 25-36.

Elie, S. D. (2006). Anthropology and post-colonial thought: the paradoxical quest for positionality. *Studies in Social and Political Thought*, *12*(1), 53-72.

Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, *24*(2), 143-188.

Fawcett, J. M., & Taylor, T. L. (2012). The control of working memory resources in intentional forgetting: Evidence from incidental probe word recognition. *Acta Psychologica*, *139*(1), 84-90.

Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PloS One, 8*(10), e77661.

Fletcher, J. (2005). Compound Rises and" Uptalk" in Spoken English. In *Ninth European Conference on Speech Communication and Technology*.

Foulkes, P. (2010). Exploring social-indexical knowledge: A long past but a short history. *Laboratory Phonology*, *1*(1), 5-39.

Foulkes, P., & Docherty, G. (2006). The social life of phonetics and phonology. *Journal of Phonetics*, *34*(4), 409-438

Foundation for Statistical Computing, Vienna, Austria. URL http://www.Rproject.com

Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review*, *118*(2), 247.

Fuchs, S., & Toda, M. (2010). Do differences in male versus female /s /reflect biological or sociophonetic factors. *Turbulent sounds: An Interdisciplinary Guide*, *21*, 281-302.

Gilbert, A. C., & Boucher, V. J. (2007). What do listeners attend to in hearing prosodic structures? investigating the human speech-parser using short-term recall. In *INTERSPEECH* (pp. 430-433).

Giles, H. (1970) Evaluative reactions to accents. *Educational Review.*, 22, 211– 27.

Giles, H. (1973). Accent mobility: A model and some data. *Anthropological Linguistics*, 87-105.

Giles, H., Coupland, J., Coupland, N., & Oatley, K. (Eds.). (1991). *Contexts of Accommodation: Developments in Applied Sociolinguistics* (Vol. 10). Cambridge University Press.

Giles, H., Coupland, J., & Coupland, N. (1991). Accommodation theory: Communication, context, and consequences. In *Contexts of Accommodation: Developments in Applied Sociolinguistics*, *1*.

Glick, P., & Fiske, S. T. (1996). The ambivalent sexism inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, *70*(3), 491.

Goldinger, S. D. (1996). Words and voices: episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(5), 1166.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological review*, *105*(2), 251. (Waterworth, J. (2003). *A Philosophical Analysis of Hope*. Springer.

Gorman, J. (1993). On Language; Like, Uptalk?. *New York Times*, *15*, 614.

Gradoville, M. (2023). The Future of Exemplar Theory. *The Handbook of Usage-Based Linguistics*, 527-544.

Gratton, C. (2016). Resisting the gender binary: The use of (ING) in the construction of non-binary transgender identities. *University of Pennsylvania Working Papers in Linguistics*, *22*(2), 7.

Green, E. (2014, January 16). A Female Senator Explains Why Uptalk Is Part of Women's 'Nature'. *The Atlantic.*

Gross, T. (2015 July 23). From upspeak to vocal fry: Are we ‘policing’ women’s voices? Radio interview with Jessica Grose and Penelope Eckert.

Guendouzi, J. (2004). 'She's very slim': talking about body-size in all-female interactions. *Journal of Pragmatics,* 36, 1635-1653.

Guy, G. R., & Vonwiller, J. (1984). The meaning of an intonation in Australian English. *Australian Journal of Linguistics*, *4*(1), 1-17.

Habasque, P. (2021). Is Creaky Voice a Valley Girl Feature? Stancetaking & Evolution of a Linguistic Stereotype. *Anglophonia. French Journal of English Linguistics*, (32).

Hachimi, A. (2016). Moroccan artists ‘blacklisted’. In *Perspectives on Arabic Linguistics XXVII: Papers from the Annual Symposium on Arabic Linguistics, Bloomington, Indiana, 2013* (Vol. 3, p. 123). John Benjamins Publishing Company.

Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. Paper presented at *Second meeting of the North American Chapter of the Association for Computational Linguistics*.

Hall, K. (2003). Exceptional speakers: Contested and problematized gender identities. In Janet Holmes and Miriam Meyerhoff (eds.) *The Handbook of Language and Gender*. Oxford , U.K. : Blackwell. 353–380.

Hall, K., & Bucholtz, M. (Eds.). (2012). *Gender articulated: Language and the Socially Constructed Self*. Routledge.

Harding, Susan. (1987) Convicted by the Holy Spirit: The Rhetoric of Fundamental Baptist Conversion. *American Ethnologist.* 14 (1): 167-181.

Hargreaves, D. A., & Tiggemann, M. (2003). Female “thin ideal” media images and boys' attitudes toward girls. *Sex Roles*, *49*, 539-544.

Harjus, J. (2017). Perceptual variety linguistics: Jerezano speakers’ concepts and perceptions of phonetic variation in western Andalusian Spanish. *Loquens*, *4*(2), e042-e042.

Harrington, J., Kleber, F., & Reubold, U. (2008). Compensation for coarticulation,/u/-fronting, and sound change in standard southern British: An acoustic and perceptual study. *The Journal of the Acoustical Society of America*, *123*(5), 2825-2835.

Hay, J. Drager, K., & Warren, P. (2010). Short-term exposure to one dialect affects processing of another. *Language and Speech 53*(4): 447-471.

Hay, J., & Drager, K. (2010). Stuffed toys and speech perception. *Linguistics 48*, 865–892.

Hay, J., Nolan, A., & Drager, K. (2006). From *fush* to *feesh*: Exemplar priming in speech perception. *The Linguistic Review, 23*(3), 351-379.

Hay, J., Warren, P., & Drager, K. (2006). Factors influencing speech perception in the context of a 2019merger-in-progress. *Journal of phonetics*, *34*(4), 458-484.

Hazenberg, E. (2019). Inferring social meaning from language variation: liminality and gender. Paper presented at United Kingdom Language Variation and Change, London, UK.

Heuer, S., & Hallowell, B. (2015). A novel eye-tracking method to assess attention allocation in individuals with and without aphasia using a dual-task paradigm. *Journal of Communication Disorders*, *55*, 15-30.

Hirschberg, J., & Ward, G. (1992). The influence of pitch range, duration, amplitude and spectral features on the interpretation of the rise-fall-rise intonation contour in English. *Journal of Phonetics*, *20*(2), 241-251.

Holden, L. (2022). Anthropologists as Experts: Cultural Expertise, Colonialism, and Positionality. *Law & Social Inquiry*, *47*(2), 669-690.

Holmes, J. (2014). Language and gender in the workplace. *The handbook of language, gender, and sexuality*, 431-451.

Horowitz, J. (27 March 2006). City girl squawk: It's like so bad – it. Really. Sucks? *New York Observer.*

Horvath, B., & Sankoff, D. (1987). Delimiting the Sydney speech community. *Language in Society*, *16*(2), 179-204.

Huskinson, T. L., & Haddock, G. (2006). Individual differences in attitude structure and the accessibility of the affective and cognitive components of attitude. *Social Cognition*, *24*(4), 453-468.

Ikeno, A., & Hansen, J. H. (2007). The effect of listener accent background on accent perception and comprehension. *EURASIP Journal on Audio, Speech, and Music Processing*, *2007*, 1-8.

Ilbury, C. (2022). U Ok Hun?: The digital commodification of white woman style. *Journal of Sociolinguistics*, *26*(4), 483-504.

Innes, B. (2007). 'Everything happened so quickly?' HRT intonation in New Zealand courtrooms. *Research on Language and Social Interaction, 40*(2-3), 227-254.

Irvine, J.T., & Gal, S. (*2000*). Language ideology and linguistic differentiation. In P. V. Kroskrity, (Ed.), *Regimes of language: Ideologies, polities, and identities*. School of American Research Advanced Seminar Series.

Jacobs-Huey, L. (2002). The natives are gazing and talking back: Reviewing the problematics of positionality, voice, and accountability among" native" anthropologists. *American Anthropologist*, *104*(3), 791-804.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, *110*(3), 306.

Jaworski, A., Coupland, N. & Galasinski, D. (2004). *Metalanguage: Social and Ideological Perspectives*. Berlin, Boston: De Gruyter Mouton.

Johnson, K., Strand, E. A., & D'Imperio, M. (1999). Auditory-visual integration of talker gender in vowel perception. *Journal of Phonetics*, *27*(4), 359-384.

Johnstone, B., & Kiesling, S. F. (2008). Indexicality and experience: Exploring the meanings of /aw/-monophthongization in Pittsburgh. *Journal of Sociolinguistics*, *12*(1), 5-33.

Johnstone, B., & Pollak, C. (2016). Mobilities, materialities, and the changing meanings of Pittsburgh speech. *Journal of English Linguistics*, *44*(3), 254-275.

Johnstone, B., Andrus, J., & Danielson, A. E. (2006). Mobility, indexicality, and the enregisterment of “Pittsburghese”. *Journal of English Linguistics*, *34*(2), 77-104.

Jonas, E., Schulz-Hardt, S., Frey, D., & Thelen, N. (2001). Confirmation bias in sequential information search after preliminary decisions: an expansion of dissonance theoretical research on selective exposure to information. *Journal of Personality and Social Psychology*, *80*(4), 557.

Jones, D. (1980). Gossip: Notes on women's oral culture. *Women's Studies International Quarterly*, *3*(2-3), 193-198.

Káldi, T., & Babarczy, A. (2021). Linguistic focus guides attention during the encoding and refreshing of Working Memory content. *Journal of Memory and Language*, *116*, 104187.

Kickham, E. A. (2015). *Purism, prescriptivism, and privilege: Choctaw language ideologies and their impact on teaching and learning.*

Kim, C., Reddy, S., Stanford, J. N., Wyschogrod, E., & Grieve, J. (2019). Bring on the crowd! Using online audio crowd-sourcing for large-scale New England dialectology and acoustic sociophonetics. *American Speech*, *94*(2), 151-194.

Kim, C., Reddy, S., Stanford, J. N., Wyschogrod, E., & Grieve, J. (2019). Bring on the crowd! Using online audio crowd-sourcing for large-scale new England dialectology and acoustic sociophonetics. *American Speech: A Quarterly of Linguistic Usage*, *94*(2), 151-194.

Kim, S. K., & Sumner, M. (2017). Beyond lexical meaning: The effect of emotional prosody on spoken word recognition. *The Journal of the Acoustical Society of America*, *142*(1), EL49-EL55.

Kimball, A. E., & Cole, J. (2016). Pitch contour shape matters in memory. In *Proceedings of the International Conference on Speech Prosody* (Vol. 8, pp. 1171-75).

Kimball, A. E., Cole, J., Dell, G. S., & Shattuck-Hufnagel, S. (2015). Categorical vs. episodic memory for pitch accents in English. *International Phonetic Association.*

King, S. (2021). Rethinking race and place: The role of persona in sound change reversal. *Journal of Sociolinguistics*, *25*(2), 159-178.

Kinzler, K. D., & DeJesus, J. M. (2013). Northern= smart and Southern= nice: The development of accent attitudes in the United States. *Quarterly Journal of Experimental Psychology*, *66*(6), 1146-1158.

Kinzler, K. D., & DeJesus, J. M. (2013). Northern= smart and Southern= nice: The development of accent attitudes in the United States. *Quarterly Journal of Experimental Psychology*, *66*(6), 1146-1158.

Kirtley, M. J. (2015). *Language, identity, and non-binary gender in Hawai'i.* PhD dissertation, University of Hawai'i at Manoa.

Klofstad, C. A., Anderson, R. C., & Peters, S. (2012). Sounds like a winner: voice pitch influences perception of leadership capacity in both men and women. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1738), 2698-2704.

Knickerbocker, H., & Altarriba, J. (2011). Bilingualism and the impact of emotion: The role of experience, memory, and sociolinguistic factors. In *Language and Bilingual Cognition* (pp. 467-492). Psychology Press.

Koops, C., Gentry, E., & Pantos, A. (2008). The effect of perceived speaker age on the perception of PIN and PEN vowels in Houston, Texas. *University of Pennsylvania Working Papers in Linguistics*, *14*(2), 12.

Kornai, A., & Kálmán, L. (1988). Hungarian sentence intonation. *Autosegmental Studies on Pitch Accent*, 183-195.

Kuhn M., Jackson S., & Cimentada, J. (2022). corrr: Correlations in R. https://github.com/tidymodels/corrr, https://corrr.tidymodels.org.

Kunda, Z. (1987). Motivated inference: Self-serving generation and evaluation of causal theories. *Journal of Personality and Social Psychology*, *53*(4), 636.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480.

Labov, W. (1963). The social motivation of a sound change. *Word*, *19*(3), 273-309.

Labov, W. (1972). Some principles of linguistic methodology. *Language in society*, *1*(1), 97-120.

Labov, W. (1990). The intersection of sex and social class in the course of linguistic change. *Language Variation andVchange*, *2*(2), 205-254.

Labov, W., Ash, S., & Boberg, C. (2008). *The atlas of North American English: Phonetics, Phonology and Sound change*. Walter de Gruyter.

Labov, W., Ash, S., Ravindranath, M., Weldon, T., Baranowski, M., & Nagy, N. (2011). Properties of the sociolinguistic monitor 1. *Journal of Sociolinguistics*, *15*(4), 431-463.

Ladd, D. Robert (1996). *Intonational Phonology*. Cambridge: Cambridge University Press.

Lakoff, R. (1973). Language and woman's place. *Language in Society*, *2*(1), 45-79.

Lakoff, R. (1975). Linguistic theory and the real world. *Language Learning*, *25*(2), 309-338.

Lawrence, D. (2017). The social perception of a sound change. *University of Pennsylvania Working Papers in Linguistics*, *23*(1), 14.

Levina, M., Waldo, C. R., & Fitzgerald, L. F. (2000). We're Here, We're Queer, We're on TV: The Effects of Visual Media on Heterosexuals' Attitudes Toward Gay Men and Lesbians. *Journal of Applied Social Psychology*, *30*(4), 738-758.

Levon, E. (2006). Hearing “gay”: Prosody, interpretation, and the affective judgments of men's speech. *American Speech*, *81*(1), 56-78.

Levon, E. (2007). Sexuality in context: Variation and the sociolinguistic perception of identity. *Language in Society*, *36*(4), 533-554.

Levon, E. (2014). Categories, stereotypes, and the linguistic perception of sexuality. *Language in Society*, *43*(5), 539-566.

Levon, E. (2016). Gender, interaction and intonational variation: The discourse functions of High Rising Terminals in London. *Journal of Sociolinguistics*, *20*(2), 133-163.

Levon, E. (2020). Same difference: the phonetic shape of High Rising Terminals in London. *English Language & Linguistics*, *24*(1), 49-73.

Levon, E., & Fox, S. (2014). Social salience and the sociolinguistic monitor: A case study of ING and TH-fronting in Britain. *Journal of English Linguistics*, *42*(3), 185-217.

Levon, E., & Ye, Y. (2020). Language, indexicality and gender ideologies: Contextual effects on the perceived credibility of women. *Gender and Language*.

Levon, E., Buchstaller, I., & Mearns, A. (2020). Towards an integrated model of perception: Linguistic architecture and the dynamics of sociolinguistic cognition. In *Advancing Socio-grammatical Variation and Change* (pp. 32-54). Routledge.

Lewandowski, T. (2012). Uptalk, Vocal Fry and, Like, Totally Slang. *Stylistyka*, (XXI), 209-219.

Linneman, T. J. (2013). Gender in Jeopardy! Intonation variation on a television game show. *Gender & Society*, *27*(1), 82-105.

Loudermilk, B. C. (2015). Implicit attitudes and the perception of sociolinguistic variation. *Responses to Language Varieties: Variability, Processes and Outcomes*, 137-156.

Luttrell, A., & Sawicki, V. (2020). Attitude strength: Distinguishing predictors versus defining features. *Social and Personality Psychology Compass*, *14*(8), e12555.

Luu, C. (2021). The Totally "Destructive" (Yet Oddly Instructive) Speech Patterns of … Young Women? *JSTOR Daily*.

Macrae, C. N., & Bodenhausen, G. V. (2001). Social cognition: Categorical person perception. *British Journal of Psychology*, *92*(1), 239-255.

Maddieson, I., & Ladefoged, P. (1993). Phonetics of partially nasal consonants. *Nasals, Nasalization, and the Velum*, p. 251-301.

Mann, V., & Repp, B. (1981). Influence of preceding fricative on stop consonant perception. *The Journal of the Acoustical Society of America*, 69, 548– 558.

Marsh, S. (28 March 2006.) 'The rise of the interrogatory statement'. *The London Times.*

MasterClass. (2021, September 22). *What Is Upspeak? 4 Tips to Break the Habit of Uptalking – 2023.*

McBride-Chang, C. (1996). Models of speech perception and phonological processing in reading. *Child Development*, *67*(4), 1836-1856.

McConnell-Ginet, S. (1983). Language, Sex and Gender: Does 'La Différence' Make a Difference?. Linguistic Society of America.

McElhinny, B. (2003). Three approaches to the study of language and gender. *American Anthropologist, 105*(4), 848 - 852

McGlothlin H., Killen M. (2010). How social experience is related to children's intergroup attitudes. *European Journal of Social Psychology*, 40, 625–634.

McGowan, K. B. (2011). *The role of socioindexical expectation in speech perception.* PhD dissertation, University of Michigan.

McGowan, K. B. (2015). Social expectation improves speech perception in noise. *Language and Speech*, *58*(4), 502-521.

McGowan, K. B., & Babel, A. M. (2019). Perceiving isn't believing: Divergence in levels of sociolinguistic awareness. *Language in Society*, *49*(2), 231-256.

McLemore, C. A. (1991). *The pragmatic interpretation of English intonation: Sorority speech*. The University of Texas at Austin.

McLennan, C. T., Luce, P. A., & Charles-Luce, J. (2003). Representation of lexical form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(4), 539.

Mendoza-Denton, N. (2014). *Homegirls: Language and cultural practice among Latina youth gangs*. John Wiley & Sons.

Mendoza-Denton, N. (2011). The semiotic hitchhiker's guide to creaky voice: Circulation and gendered hardcore in a Chicana/o gang persona. *Journal of Linguistic Anthropology*, *21*(2), 261-280.

Meyerhoff, M., & Ehrlich, S. (2019). Language, gender, and sexuality. *Annual Review of Linguistics*, *5*, 455-475.

Hess, S. (2015, May 6). *Millennials: Who They Are and Why We Hate Them by Scott Hess*. Millennials in the Workplace.

Mills, S. (2002). *Feminist stylistics*. Routledge.

Milroy, J., & Milroy, L. (1985). Linguistic change, social network and speaker innovation1. *Journal of Linguistics*, *21*(2), 339-384.

Milroy, L. (2004). Language ideologies and linguistic change. *Sociolinguistic Variation: Critical Reflections*, 161-177.

Morgan, E. (14 July 2015). 99% Invsible Podcast's brilliant response to criticims of women's voices and what we really need to be talking about. *Medium.*

Mullenix, J. (2005). The perceptual processing of vocal emotion. *Psychology of Moods. Nova Science Publishers, Hauppage, NY*, 123-139.

Mullenix, J. W., Johnson, K. A., Topcu-Durgun, M., & Farnsworth, L. M. (1995). The perceptual representation of voice gender. *Journal of the Acoustical Society of America, 98*(6), 3080-3095.

Munson, B. (2011). The influence of actual and imputed talker gender on fricative perception, revisited *The Journal of the Acoustical Society of America*, *130*(5), 2631-2634.

Munson, B., Ryherd, K., & Kemper, S. (2017). Implicit and explicit gender priming in English lingual sibilant fricative perception. *Linguistics*, *55*(5), 1073-1107.

Nagy, N., Chociej, J., & Hoffman, M. F. (2014). Analyzing Ethnic Orientation in the quantitative sociolinguistic paradigm. *Language & Communication*, *35*, 9-26.

Nakamura, M. (2014). *Gender, language and ideology: A genealogy of Japanese women's language* (Vol. 58). John Benjamins Publishing Company.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175-220.

Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology*, *18*(1), 62-85.

Nord, W. R. (1969). Social exchange theory: an integrative approach to social conformity. *Psychological Bulletin*, *71*(3), 174.

Norman, D. A. (1968). Toward a theory of memory and attention. *Psychological Review*, *75*(6), 522.

Nunberg, G. (2016, May 24). Irked By The Way Millennials Speak? "I Feel Like" It's Time To Loosen Up. *WEMU-FM*.

Nycum, R. (2018). In Defense of Valley Girl English. *The Compass* (Vol. 1, No. 5, p. 4).

Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception & psychophysics*, *60*(3), 355-376.

Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science*, *5*(1), 42-46.

O'Grady, H. (2005). *Woman's relationship with herself: Gender, Foucault and therapy*. Routledge.

Oakley, T. (2009). *From attention to meaning: Explorations in semiotics, linguistics, and rhetoric* (Vol. 8). Peter Lang.

Ochs, E. (1993). Constructing social identity: A language socialization perspective. *Research on language and social interaction*, *26*(3), 287-306.

Okely, J. (2012). Confronting positionality. *Teaching Anthropology*, *2*(1).

Olson, J. M., Vernon, P. A., Harris, J. A., & Jang, K. L. (2001). The heritability of attitudes: a study of twins. *Journal of Personality and Social Psychology*, *80*(6), 845.

Oosterhoff, B., Shook, N. J., & Ford, C. (2018). Is that disgust I see? Political ideology and biased visual attention. *Behavioural Brain Research*, *336*, 227-235.

Ortega, M. (2006). Being lovingly, knowingly ignorant: White feminism and women of color. *Hypatia*, *21*(3), 56-74.

Oswald, M. E., & Grosjean, S. (2004). Confirmation bias. *Cognitive Illusions: A Handbook on Fallacies and Biases in Thinking, Judgement and Memory*, *79*.

Owren, M. J., Berkowitz, M., & Bachorowski, J. A. (2007). Listeners judge talker sex more efficiently from male than from female vowels. *Perception & Psychophysics*, *69*(6), 930-941.

Palmeri, T. J., Goldinger, S. D., & Pisoni, D. B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(2), 309.

Parnell-Mooney, J. (2019). *Tran/s/gender: assessing the effects of the social construction of gender on speech: a focus on transgender/s/realisations.* PhD dissertation, University of Glasgow.

Phillips, W., & Boroditsky, L. (2003). Can quirks of grammar affect the way you think? Grammatical gender and object concepts. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 25, No. 25).

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. *Typological studies in Language*, *45*, 137-158.

Pierrehumbert, J. B. (2016). Phonological representation: Beyond abstract versus episodic. *Annual Review of Linguistics*, *2*, 33-52.

Pisoni, D. B., & Levi, S. V. (2007). Representations and representational specificity in speech perception and spoken word recognition.

Podesva, R. J. (2011). Salience and the social meaning of declarative contours: Three case studies of gay professionals. *Journal of English Linguistics*, *39*(3), 233-264.

Podesva, R. J., & Kajino, S. (2014). Sociophonetics, gender, and sexuality. *The Handbook of Language, Gender, and Sexuality*, 103-122.

Podesva, R. J., & Szakay, A. (2013). Gender differences in the acoustic realization of creaky voice: Evidence from conversational data collected in Northern California. *The Journal of the Acoustical Society of America*, *134*(5_Supplement), 4238-4238.

Podesva, R. J., Callier, P., & Jamsu, J. (2012). Recency, resonance, and the structuring of phonological style in political speeches. *JM Hernández-Campoy y JA Cutillas-Espinosa,(eds.)*, 101-117.

Prechtel, C., & Clopper, C. G. (2016). Uptalk in Midwestern American English. In *Proceedings of the 8th International Conference on Speech Prosody. Edited by Jon Barnes, Alejna Burgos, Stefanie Shattuck-Hufnagel and Nanette Veilleux. International Speech Communication Association (ISCA) Online Archive* (pp. 133-37).

Quenqua, D. (2012). They're, like, way ahead of the linguistic currrrve. *The New York Times*, *28*.

R Core Team. 2012. Team R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Ranganath, C., & Rainer, G. (2003). Neural mechanisms for detecting and remembering novel events. *Nature Reviews Neuroscience*, *4*(3), 193-202.

Reno, R. R., Cialdini, R. B., & Kallgren, C. A. (1993). The trans-situational influence of social norms. *Journal of Personality and Social Psychology*, *64*(1), 104.

Rickford, J. R., & King, S. (2016). Language and linguistics on trial: Hearing Rachel Jeantel (and other vernacular speakers) in the courtroom and beyond. *Language*, 948-988.

Riley, E., Okabe, H., Germine, L., Wilmer, J., Esterman, M., & DeGutis, J. (2016). Gender differences in sustained attentional control relate to gender inequality across countries. *PloS One*, *11*(11), e0165100.

Ritchart, A., & Arvaniti, A. (2014). The form and use of uptalk in Southern Californian English.

Robinson, P. M. (Ed.). (2003). *Time series with long memory*. Advanced Texts in Econometrics. Oxford University Press.

Roettger, T. B., & Rimland, K. (2020). Listeners' adaptation to unreliable intonation is speaker-sensitive. *Cognition*, *204*, 104372.

Rogers, H. (2000). *The sounds of language: An introduction to phonetics.* Harlow, UK: Pearson Education.

Romm, C. (2017, June 2). The Real Purpose of Uptalk Is to Get You to Shut Up and Listen. *The Cut*.

Rosa, J. D. (2016). Standardization, racialization, languagelessness: Raciolinguistic ideologies across communicative contexts. *Journal of Linguistic Anthropology*, *26*(2), 162-183.

Rubin, D. L. (1992). Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research in Higher Education*, *33*, 511-531.

Sankoff, D., & Laberge, S. (1978). The linguistic market and the statistical explanation of variability. *Linguistic Variation: Models and Methods*, *239*, 250.

Schilling-Estes, N. (1998). Investigating “self-conscious” speech: The performance register in Ocracoke English. *Language in Society*, *27*(1), 53-83.

Schleef, E. (2005). Gender, power, discipline and context: On the sociolinguistic variation of okay, right, like, and you know in English Academic Discourse. In *Twelfth Annual Symposium about Language and Society - Austin, TX* (Vol. 48, pp. 177-186).

Schmid, M. S., Berends, S. M., Bergmann, C., Brouwer, S. M., Meulman, N., Seton, B. J., & Sprenger, S. A. (2016). Eye-tracking and the visual world paradigm. *Designing Research on Bilingual Development: Behavioral and Neurolinguistic Experiments*, 55-80.

Schmid, M., & Bradley, E. (2019). Vocal pitch and intonation characteristics of those who are gender non-binary. In *Proceedings of the 19th International Conference of Phonetic Sciences* (pp. 2685-2689).

Seitz-Brown, M. (2014, December 16). Young Women Shouldn't Have to Talk Like Men to Be Taken Seriously. *Slate Magazine*. https://slate.com/human-interest/2014/12/uptalk-is-ok-young-women-shouldn-t-have-to-talk-like-men-to-be-taken-seriously.html

Severijnen, G. G., Bosker, H. R., Piai, V., & McQueen, J. M. (2021). Listeners track talker-specific prosody to deal with talker-variability. *Brain Research*, *1769*, 147605.

Sharma, D., & McCarthy, K. (2018). Attentional load and style control. *University of Pennsylvania Working Papers in Linguistics*, *24*(2), 15.

Shokeir, V. (2008). Evidence for the stable use of uptalk in South Ontario English. *University of Pennsylvania Working Papers in Linguistics*, *14*(2), 4.

Sigmon, S.T., Stanton, A.L. & Snyder, C.R. (1995). Gender differences in coping: A further test of socialization and role constraint theories. *Sex Roles* 33, 565–587.

Silverstein, M. (2003). Indexical order and the dialectics of sociolinguistic life. *Language & Communication*, *23*(3-4), 193-229.

Simmons, J. P., & Prentice, D. A. (2006). Pay attention! Attention to the primes increases attitude assessment accuracy. *Journal of Experimental Social Psychology*, *42*(6), 784-791.

Slobe, T. (2018). Style, stance, and social meaning in mock white girl. *Language in Society*, *47*(4), 541-567.

Slobe, T. (2016). This American Creak: Metaphors of Virus, Infection, and Contagion in Girls' Social Networks. Paper presented at the American Anthropological Association in Minneapolis, MN.

Smith, D. R., & Patterson, R. D. (2005). The interaction of glottal-pulse rate and vocal-tract length in judgements of speaker size, sex, and age. *The Journal of the Acoustical Society of America*, *118*(5), 3177-3186.

Smith, D. R., Walters, T. C., & Patterson, R. D. (2007). Discrimination of speaker sex and size when glottal-pulse rate and vocal-tract length are controlled. *The Journal of the Acoustical Society of America*, *122*(6), 3628-3639.

Smith, D.R., Patterson, R.D., Turner, Kawahara, H., Irino, T. (2005). The processing and perception of size information in speech sounds. *Journal of the Acoustical Society of America 2005,* 117, 305–318.

Song, Y. J., Clopper, C. G., & Wagner, L. (2022). Children's Use of Uptalk in Narratives. *Proc. Speech Prosody 2022*, 32-36.

Squire, L. R. (2004). Memory systems of the brain: a brief history and current perspective. *Neurobiology of Learning and Memory*, *82*(3), 171-177.

Squires, L. (2013). It don't go both ways: Limited bidirectionality in sociolinguistic perception. *Journal of Sociolinguistics*, *17*(2), 200-237.

Squires, L. (2014). Processing, evaluation, knowledge: Testing the perception of English subject–verb agreement variation. *Journal of English Linguistics*, *42*(2), 144-172.

Squires, L. (2016). Processing grammatical differences: Perceiving versus noticing. *Awareness and Control in Sociolinguistic Research*, 80-103.

Staum Casasanto, L. (2009). Experimental investigations of sociolinguistic knowledge. Unpublished PhD dissertation. Palo Alto, California: Stanford University.

Staum Casasanto, L. (2009). How Do Listeners Represent Sociolinguistic Knowledge?. In *Proceedings of the Annual Meeting of the Cognitive Science Society, 31*.

Stecker, A. (2020). Investigations of the sociolinguistic monitor and perceived gender identity. *University of Pennsylvania Working Papers in Linguistics*, *26*(2), 14.

Stecker, A., & D'Onofrio, A. (2022). Variation in Evaluations of Gendered Voices: Individual Speakers Condition the Variant Frequency Effect. *Journal of English Linguistics*, *50*(3), 281-314.

Steele, A. J. (2019). *Non-binary speech, race, and non-normative gender: Sociolinguistic style beyond the binary*. Doctoral dissertation, The Ohio State University.

Stevens, J.P. (1992). Applied multivariate statistics for the social sciences (2nd edition). Hillsdale, NJ: Erlbaum.

Strand, E. A. (1999) "Uncovering the role of gender stereotypes in speech perception." *Journal of Language and Social Psychology* 18.1 (1999): 86-100.

Sumner, M. (2015). The social weight of spoken words. *Trends in Cognitive Sciences*, *19*(5), 238-239.

Sumner, M., & Kataoka, R. (2013). Effects of phonetically-cued talker variation on semantic encoding. *The Journal of the Acoustical Society of America*, *134*(6), EL485-EL491.

Sumner, M., & Samuel, A. G. (2005). Perception and representation of regular variation: The case of final/t. *Journal of Memory and Language*, *52*(3), 322-338.

Sumner, M., Kim, S. K., King, E., & McGowan, K. B. (2014). The socially weighted encoding of spoken words: A dual-route approach to speech perception. *Frontiers in Psychology*, *4*, 1015.

Szakay, A., Babel, M., & King, J. (2012). Sociophonetic markers facilitate translation priming: Maori English GOAT–A different kind of animal. *University of Pennsylvania Working Papers in Linguistics*, *18*(2), 16.

Tajfel, H. (2010). Social stereotypes and social groups. In T. Postmes & N. R. Branscombe (Eds.), *Rediscovering Social Identity* (pp. 191–206). Psychology Press.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: a social psychological perspective on mental health. *Psychological Bulletin*, *103*(2), 193.

Thibault, Pierrette. 1983. *Équivalence et grammaticalisation*. Montreal: University of Montreal dissertation.

Thomas, E. R. (2013). Sociophonetics. *The Handbook of Language Variation and Change*, 108-127.

Thomson, S. (2018, September 21). Why the Confidence Gap Is a Myth. *The Atlantic*.

Tobin, E., & Benders, T. (2018). Interpretations of Uptalk in Australian English: Low confidence, unfinished speech, and variability within and between listeners. In *Australasian International Conference on Speech Science and Technology (17th: 2018)*, p. 9-12.

Tooley, K. M., Konopka, A. E., & Watson, D. G. (2014). Can intonational phrase structure be primed (like syntactic structure)?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(2), 348.

Tripp, A., & Munson, B. (2022). Perceiving gender while perceiving language: Integrating psycholinguistics and gender theory. *Wiley Interdisciplinary Reviews: Cognitive Science*, *13*(2), e1583.

Tyler, J. C. (2015). Expanding and mapping the indexical field: Rising pitch, the uptalk stereotype, and perceptual variation. *Journal of English Linguistics*, *43*(4), 284-310.

Vaissière, J. (1995). Phonetic explanations for cross-linguistic prosodic similarities. *Phonetica*, *52*(3), 123-130.

Van Berkum, J. J., Brown, C. M., & Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of memory and language*, *41*(2), 147-182.

Vaughn, C., & Kendall, T. (2018). Listener sensitivity to probabilistic conditioning of sociolinguistic variables: The case of (ING). *Journal of Memory and Language*, *103*, 58-73.

Villarreal, Daniel. (2016). The construction of social meaning: A matched-guise investigation of the California Vowel Shift. PhD dissertation, University of California, Davis.

Vinacke, W. E. (1957). Stereotypes as social concepts. *The Journal of Social Psychology*, *46*(2), 229-243.

Vornik, L., Sharman, S., & Garry, M. (2003). The power of the spoken word: Sociolinguistic cues influence the misinformation effect. *Memory*, *11*(1), 101-109.

Wagner, S. E., & Hesson, A. (2014). Individual sensitivity to the frequency of socially meaningful linguistic cues affects language attitudes. *Journal of Language and Social Psychology*, *33*(6), 651-666.

Walker, A., García, C., Cortés, Y., & Campbell-Kibler, K. (2014). Comparing social meanings across listener and speaker groups: The indexical field of Spanish/s. *Language Variation and Change*, *26*(2), 169-189.

Warren, P. (2016). *Uptalk: The phenomenon of rising intonation*. Cambridge University Press.

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology*, *12*(3), 129-140.

Waterworth, J. & Riva, G. (2014). The Designed Presence of the Individual. *Feeling Present in the Physical World and in Computer-Mediated Environments*, 57-77.

Wilson Black, J., Brand, J., Hay, J., & Clark, L. (2023). Using principal component analysis to explore co-variation of vowels. *Language and Linguistics Compass*, *17*(1), e12479.

Winn, M. B., & Moore, A. N. (2020). Perceptual weighting of acoustic cues for accommodating gender-related talker differences heard by listeners with normal hearing and with cochlear implants. *The Journal of the Acoustical Society of America*, *148*(2), 496-510.

Winn, M. B., Tripp, A., & Munson, B. (2022). A critique and call for action, in response to sexist commentary about vocal fry. *Perspectives of the ASHA Special Interest Groups*, *7*(6), 1903-1907.

Wodak, R., & Benke, G. (2017). Gender as a sociolinguistic variable: New perspectives on variation studies. *The handbook of sociolinguistics*, 127-150.

Wodak, Ruth. (2000). Gender as Sociolinguistics Variable. Florian Coulmas (Ed). *The Handbook of Sociolinguistics.* (Hlm 129-150).

Wolf, N. (2017, November 29). Young women, give up the vocal fry and reclaim your strong female voice. *The Guardian*. https://www.theguardian.com/commentisfree/2015/jul/24/vocal-fry-strong-female-voice

Wolfram, W. A. (1969). *A sociolinguistic description of Detroit Negro speech. Urban Language Series, No. 5.*

Wollum, E. (2019). The Uptalk Downgrade: Comparing age-and gender-based perceptions of uptalk in four highly-skilled professions. Master's thesis, Victoria University of Wellington.

Yi, H. G., Phelps, J. E., Smiljanic, R., & Chandrasekaran, B. (2013). Reduced efficiency of audiovisual integration for nonnative speech. *The Journal of the Acoustical Society of America*, *134*(5), EL387-EL393.

Yoodli (2022, December 19). What Is Uptalk? Why, When, & How to Stop Doing It.

Yuasa, I. P. (2010). Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women?. *American Speech*, *85*(3), 315-337.

Zanna, M. P., & Rempel, J. K. (1988). Attitudes: A new look at an old concept. In D. Bar-Tal & A. W. Kruglanski (Eds.), *The social psychology of knowledge* (pp.315-334). Cambridge, England: Cambridge University Press.

Zarcone, A., Van Schijndel, M., Vogels, J., & Demberg, V. (2016). Salience and attention in surprisal-based accounts of language processing. *Frontiers in Psychology*, *7*, 844.

Zimman, L. (2012). Voices in transition: Testosterone, transmasculinity, and the gendered voice among female-to-male transgender people. Doctoral dissertation, University of Colorado at Boulder.

Zimman, L. (2013). Hegemonic masculinity and the variability of gay-sounding speech: The perceived sexuality of transgender men. *Journal of Language & Sexuality 2*(1):5-43.

Zimman, L. (2017). Gender as stylistic bricolage: Transmasculine voices and the relationship between fundamental frequency and/s. *Language in Society*, *46*(3), 339-370.

Zimman, L. (2017). Variability in /s/ among transgender speakers: Evidence for a socially-grounded account of gender and sibilants. *Linguistics* 55(5): 993-1019.

Zimman, L. (2018). Transgender voices: Insights on identity, embodiment, and the gender of the voice. *Language and Linguistics Compass 12*(7):e12284.

Zimman, L. (2019). Trans self-identification and the language of neoliberal selfhood: Agency, power, and the limits of monologic discourse. *International Journal of the Sociology of Language*, *2019*(256), 147-175.

Zormpa, E., Brehm, L. E., Hoedemaker, R. S., & Meyer, A. S. (2019). The production effect and the generation effect improve memory in picture naming. *Memory*, *27*(3), 340-352.

Appendix

Appendix. A. Stimulus passage with original (old) utterance contours, divided across four experimental blocks. Studies 1 and 3 used all blocks; Study 2 used Block 1 with all falls and Block 2 with the original contour scheme and ending frame.

Sentence #	Block	Sentence	contour
intro frame	Block 1	Hey there! Sam here, and I'm back with another segment of Make the Leap –the podcast that dives into all kinds of facts about our favorite amphibians. Today, we're sprouting legs and leaping into today's animal topic: frogs!	filler
1		Some quick backstory on why we chose to take a peak at FROGs today	fall
2		I was cleaning out the closet in my hallway	rise
3		I found the terrarium that my brother and I shared	fall
		when we caught frogs in the backyard.	filler
5		Most of them would escape	rise
6		But I remember we actually kept some for a while	rise
7		So it's a fun memory nonetheless.	fall
11		The world's largest frog is the goliath frog of West Africa	rise
		Which can grow to 15 inches	filler
13		And this frog can also weigh up to seven pounds.	fall
14	Block 2	One of the smallest frogs on the planet is the Cuban tree frog,	rise
		Where some of the largest grow to just half an inch long.	filler
16		We still don't really know how long frogs live in the wild	rise
17		But frogs in captivity have been known to live more than 20 years.	fall
18		There are more than 6,000 species of frogs worldwide.	fall
19		And scientists are still on the search for new ones.	fall
20		However, frogs are actually slowly going extinct	rise
		Apparently amphibians have a high extinction rate relative to other kinds of animals,	filler
22		And more than 200 species of frogs have gone extinct since the 1970s	fall
25		Many frogs can change color out of adaption to survive	rise
26	Block 3	But Most cannot completely change their color,	fall
27		Instead they change the shade of their basic color.	fall
30		They also change color to help control their body temperature	rise
		As some colors absorb more or less light.	filler
		Thinking about reproduction,	filler

33		Sometimes frogs produce ribbit sounds	rise
		When they are trying to attract a mate	filler
35		This brings me back to our terrarium	rise
36		At one point we actually had frog eggs growing	fall
37		We started hatching these frogs eggs eventually	fall
38		But only because we carried the terrarium outside to the back of the yard	rise
39		And our parents must have been so mad at us.	fall
40		Most frogs lay their eggs in water,	rise
41		But you can find some exceptions.	fall
42		Frog eggs don't have shells,	rise
		So they need some kind of moisture to keep them from drying out until they hatch.	filler
44	Block 4	Some frogs have come up with amazing ways to maintain moisture with their eggs	rise
		Besides just laying them directly in water.	fall
46		Some frogs can lay eggs under leaves even above water in rainforests where it’s damp	rise
47		And when they hatch, the tadpoles fall into the water	fall
48		And there they can develop even further.	fall
end		Well, that completes our fast facts and fun times with frogs! Tune in next week for our next amphibian segment.	filler

Appendix. B. Norming 2 speaker social and demographic evaluations and uptalk perception question responses.

Appendix. B. 1. Percentages of perceived gender and average ratings of gender prototypicality (1: "not at all prototypical" to 5: "very prototypical") by speaker.

Gender						Gender prototypicality		
Speaker	Gender	Another gender	Man	Woman		Male speaker prototypicality	Female speaker prototypicality	Non-binary speaker prototypicality
AS	FF	0%	0%	100%		1.067	4.7	2.333
AV	M	0%	100%	0%		4.7	1.167	2.2
DD	F	0%	0%	100%		1.1	4.57	2.533
ES	F	0%	0%	100%		1.1	4.9	2.4
GF	M	0%	100%	0%		4.6	1.1	2.5
JO	F	0%	0%	100%		1.033	4.867	2.267
JS	F	3%	0%	97%		1.067	4.767	2.467
NZ	M	0%	100%	0%		4.633	1.1	2.6
PS	M	0%	100%	0%		4.533	1.2	2.43
TS	M	0%	97%	0%		4.6	1.13	2.6

Appendix. B. 2. Percentages of perceived age by speaker.

Age				
Speaker	18-30 years old	30-40 years old	40-50 years old	Younger than 18
AS	83%	13%	0%	3%
AV	70%	30%	0%	0%
DD	83%	10%	0%	7%
ES	87%	7%	0%	7%
GF	87%	10%	0%	3%
JO	97%	3%	0%	0%
JS	80%	20%	0%	0%
NZ	87%	13%	0%	0%
PS	50%	47%	3%	0%
TS	73%	27%	0%	0%

Appendix. B. 3. Percentages of perceived origin of geographic region by speaker.

Region					
Speaker	Midwest	Northeast	Other	South	West
AS	37%	30%	3%	3%	27%
AV	40%	33%	0%	7%	20%
DD	33%	33%	0%	3%	30%
ES	30%	23%	0%	10%	37%
GF	30%	23%	0%	10%	37%
JO	40%	30%	0%	3%	27%
JS	23%	43%	0%	3%	30%
NZ	50%	30%	0%	3%	17%
PS	43%	30%	0%	3%	23%
TS	40%	23%	0%	10%	27%

Appendix. B. 4. Percentages of perceived racialized identity by speaker.

Racialized identity									
Speaker	Asian	Asian, Latinx, White	Asian, White	Black or African American	Latinx	Latinx, Other	Latinx, White	Other	White
AS	7%	3%	0%	0%	7%	0%	3%	0%	80%
AV	0%	3%	0%	0%	3%	0%	0%	0%	93%
DD	3%	0%	3%	0%	0%	0%	0%	3%	90%
ES	0%	0%	0%	7%	0%	0%	0%	0%	93%
GF	3%	0%	0%	0%	0%	3%	3%	3%	87%
JO	0%	0%	3%	3%	3%	0%	0%	0%	90%
JS	0%	0%	0%	0%	3%	0%	0%	3%	93%
NZ	0%	3%	3%	3%	3%	0%	0%	0%	87%
PS	3%	0%	0%	0%	0%	0%	0%	3%	93%
TS	3%	0%	3%	3%	0%	0%	0%	0%	90%

Appendix. B. 5. Perceived US English nativeness by speaker.

Speaker	Yes, native speaker	No, non-native speaker	Not sure
AS	93%	0%	7%
AV	97%	0%	3%
DD	97%	0%	3%
ES	97%	0%	3%
GF	93%	0%	7%
JO	97%	3%	0%
JS	97%	0%	3%
NZ	100%	0%	0%
PS	96.7%	0%	3%
TS	96.7%	0%	3%

Appendix. B. 6. Self-identified demographic information for speakers serving as stimuli.

Speaker	Gender	Age	Racialized identity	Region of origin
AS	F	25	White	Midwest
DD	F	26	Multiracial	Midwest
ES	F	24	White	Midwest
JO	F	27	White	Midwest
JS	F	27	White	Midwest
AV	M	26	White	Pacific Northwest
GF	M	30	White	Midwest
NZ	M	26	White	Midwest
PS	M	29	White	Midwest
TS	M	25	White	South

Appendix. B. 7. Average social evaluation ratings for each speaker.

Speaker	Gender	Professional	Friendly	Engaging	Attractive	Nerdy	Intelligent	Enthusiastic	Speech rate
AS	F	3.8	3.667	3.567	3.5	2.5	3.767	3.4	3.1
AV	M	3.533	3.80	3.333	3.233	3.2	3.5	3.133	2.8
DD	F	3.033	2.933	2.467	3.133	2.6	3.3	2.2	2.3
ES	F	3.9	4.467	4.3	3.767	2.8	3.867	4.367	3.333
GF	M	3.367	3.567	2.767	3.033	3.033	3.433	3.033	2.8
JO	F	3.7	4.233	3.867	3.5	2.667	3.733	3.967	3.367
JS	F	3.833	4.167	4.1	3.667	2.667	3.833	4	4
NZ	M	3.5	3.567	3.2	2.967	2.867	3.3	2.867	2.867
PS	M	3.567	3.767	3.333	3.1	2.733	3.467	3.367	3.033
TS	M	3.767	3.733	3.4	3.1	3.067	3.667	2.967	2.867

Appendix. B. 8. Average uptalk perceivability (1: "I didn't notice it all" to 5: "I noticed immediately and constantly") and percentages of classification of novel uptalk utterance ("Does this sentence contain uptalk?") by speaker.

Speaker	Gender	Average uptalk perceivability	Yes: Statement contained uptalk	No: Statement did not contain uptalk
AV	M	3.13	93%	7%
AS	F	2.7	37%	47%
DD	F	2.33	53%	33%
ES	F	2.77	77%	20%
GF	M	2.73	83%	13%
JO	F	2.9	73%	17%
JS	F	2.87	70%	20%
NZ	M	2.97	73%	17%
PS	M	3.07	83%	7%
TS	M	2.4	47%	40%

Appendix. C. Norming 3 accuracy for classification of each utterance as containing "uptalk" or a "falling "tone."

Sentence #	contour	accuracy		Sentence #	contour	accuracy
1	fall	0.847		1	rise	0.879
2	fall	0.879		2	rise	0.847
3	fall	0.883		3	rise	0.814
5	fall	0.864		5	rise	0.729
6	fall	0.915		6	rise	0.780
7	fall	0.879		7	rise	0.847
8	fall	0.864		8	rise	0.862
10	fall	0.828		10	rise	0.787
11	fall	0.871		11	rise	0.915
13	fall	0.879		13	rise	0.852
14	fall	0.917		14	rise	0.931
16	fall	0.931		16	rise	0.797
17	fall	0.900		17	rise	0.915
18	fall	0.862		18	rise	0.831
19	fall	0.881		19	rise	0.862
20	fall	0.862		20	rise	0.898
22	fall	0.915		22	rise	0.931
23	fall	0.897		23	rise	0.864
24	fall	0.847		24	rise	0.914
25	fall	0.897		25	rise	0.903
26	fall	0.886		26	rise	0.816
27	fall	0.915		27	rise	0.881
28	fall	0.810		28	rise	0.881
29	fall	0.852		29	rise	0.897
30	fall	0.914		30	rise	0.831
33	fall	0.881		33	rise	0.897
35	fall	0.914		35	rise	0.780
36	fall	0.885		36	rise	0.931
37	fall	0.850		37	rise	0.900
38	fall	0.932		38	rise	0.914
39	fall	0.862		39	rise	0.763
40	fall	0.902		40	rise	0.931
41	fall	0.879		41	rise	0.847
42	fall	0.915		42	rise	0.931
44	fall	0.817		44	rise	0.831
46	fall	0.898		46	rise	0.862
47	fall	0.897		47	rise	0.867
48	fall	0.935		48	rise	0.879
49	fall	0.914		49	rise	0.915

50	fall	0.790		50	rise	0.914

Appendix. D. F0 characteristics predicted by speaker and speaker gender. Individual speaker coefficients indicate difference from grand mean.

Relative rise excursions (ERB)				
	Estimate	Std. Error	T-value	P-value
(Intercept)	0.657	0.025	26.011	< 0.0001 ***
AV	-0.218	0.037	-5.943	< 0.0001 ***
ES	0.003	0.038	0.082	0.935
GF	0.099	0.037	2.709	0.008 *
JO	0.076	0.037	2.044	0.043 *
JS	0.035	0.037	0.936	0.351
NZ	0.007	0.035	0.181	0.858
Gender (=man)	-0.077	0.036	-2.121	0.036 *
Absolute rise excursion (ERB)				
	Estimate	Std. Error	T-value	P-value
(Intercept)	2.726	0.116	23.513	<0.0001 ***
AV	-0.129	0.177	-0.726	0.469
ES	0.632	0.185	3.420	0.001 **
GF	-0.848	0.177	-4.788	<0.0001 ***
JO	0.350	0.180	1.949	0.053
JS	0.042	0.180	0.232	0.817
NZ	0.041	0.169	-0.28368	0.7786
Gender (=man)	-0.679	0.1658	-4.098	<0.0001 ***
Slope (ERB/ms)				
	Estimate	Std. Error	T-value	P-value
(Intercept)	0.013	0.001	17.143	< 0.0001 ***
AV	0.000	0.001	-0.214	0.831
ES	0.003	0.001	2.124	0.035 *
GF	-0.006	0.001	-4.216	< 0.0001 ***
JO	0.005	0.001	3.434	0.001 **
JS	-0.001	0.001	-0.410	0.683
NZ	-0.001	0.000	-0.906	0.372
Gender (=man)	-0.005	0.001	-3.762	0.0002 ***

F0 Range (ERB)				
	Estimate	Std. Error	T-value	P-value
(Intercept)	4.446	0.129	34.445	< 0.0001 ***
AV	1.744	0.198	8.808	< 0.0001 ***
ES	0.702	0.206	3.401	0.001 **
GF	-1.802	0.198	-9.114	< 0.0001 ***
JO	-0.067	0.200	-0.337	0.737
JS	-0.496	0.200	-2.474	0.014 *
NZ	-0.090	0.189	0.317	0.754
Gender (=man)	-0.081	0.249	-0.326	0.745
F0 space: rise pitch slope (ERB/ms) – fall pitch slope (ERB/ms)				
	Estimate	Std. Error	T-value	P-value
(Intercept)	139.821	3.998	34.972	< 0.0001 ***
AV	-26.456	5.828	-4.54	< 0.0001 ***
ES	60.051	5.837	10.288	< 0.0001 ***
GF	-62.948	5.827	-10.803	< 0.0001 ***
JO	44.871	5.825	7.703	< 0.0001 ***
JS	18.96	5.823	3.256	0.00124 **
NZ	-34.602	5.825	10.578	< 0.0001 ***
Gender (=man)	-82.555	5.463	-15.11	< 0.0001 ***

Appendix. E. Differences in F0 space between falling and rising versions of each utterance.

Falling utterances					Rising utterances				
Sentence	Onset F0 (Hz)	Terminal F0 (Hz)	onset - terminal difference		Sentence	Onset F0 (Hz)	Terminal F0 (Hz)	onset - terminal difference	fall and rise F0 difference
1	225	91	134		1	179	255	-76	210
2	163	108	55		2	166	234	-68	129
3	172	109	63		3	168	232	-64	127
5	190	139	50		5	173	252	-79	129
6	188	118	71		6	175	252	-77	148
7	186	108	79		7	186	270	-84	162
11	178	141	37		11	173	255	-82	119
13	166	145	21		13	169	242	-73	95
14	179	114	65		14	181	274	-94	159
16	185	114	71		16	185	245	-60	130
17	178	114	64		17	180	237	-57	121
18	194	123	71		18	169	248	-79	150
19	187	115	72		19	156	224	-67	144
20	184	111	73		20	183	271	-88	161
22	174	117	57		22	174	266	-92	150
25	190	121	70		25	169	229	-60	130
26	184	97	87		26	188	270	-82	161
27	176	108	68		27	178	235	-58	133
30	149	104	44		30	168	259	-91	119
33	203	152	52		33	185	251	-66	111
35	206	105	101		35	195	272	-77	184
36	176	139	38		36	160	250	-90	109
37	192	143	49		37	190	266	-75	118
38	169	121	48		38	180	251	-71	122
39	139	107	32		39	179	238	-60	97
40	185	106	79		40	206	289	-82	149
41	209	119	90		41	180	273	-93	173
42	189	133	56		42	177	242	-65	120
44	196	114	82		44	198	251	-53	135
46	206	101	105		46	175	233	-57	153
47	163	111	51		47	184	290	-106	134
48	162	119	43		48	174	254	-80	125

Appendix. F. Cloze task items in each block, target items emboldened. Studies 1 and 3 used all blocks; Study 2 used Blocks 1 and 2.

Block	Trial
	I was cleaning out the closet in my **hallway**
	I found the terrarium that my brother and I **shared**
	but I remember we **actually** kept some for a while
Block 1	so it's a **fun** memory nonetheless.
	But frogs in **captivity** have been known to live more than 20 years
	And scientists are still on the **search** for new ones
	However, frogs are **actually** slowly going extinct
Block 2	Many frogs can change color out of **adaption** to survive
	instead they change the shade of their **basic** color
	They also change color to help **control** their body temperature
	This brings me back to our **terrarium**
Block 3	We started hatching these frogs eggs **eventually**
	but only because we **carried** the terrarium outside to the back of the yard
	Most frogs lay their eggs in **water**
	and when they hatch the tadpoles **fall** into the water
Block 4	and there they can **develop** even further

Appendix. G. Study 3 Tukey HSD tests for significant differences in "old" responses between metalinguistic prime conditions.

Appendix. G. 1. Tukey HSD comparisons of average hit rates between prime conditions and utterance contour types.

Comparison	Difference	Lower CI	Upper CI	Adjusted p-value
equal prime:fall - baseline:fall	0.013	-0.115	0.140	1.000
female bias prime:fall - baseline:fall	-0.111	-0.223	0.001	0.049 *
male bias prime:fall - baseline:fall	-0.043	-0.162	0.076	0.959
baseline:rise - baseline:fall	-0.040	-0.151	0.071	0.956
equal prime:rise - baseline:fall	-0.051	-0.179	0.076	0.927
female bias prime:rise - baseline:fall	-0.026	-0.138	0.086	0.997
male bias prime:rise - baseline:fall	-0.112	-0.231	0.007	0.084
female bias prime:fall - equal prime:fall	-0.124	-0.252	0.004	0.066
male bias prime:fall - equal prime:fall	-0.055	-0.190	0.079	0.917
baseline:rise - equal prime:fall	-0.053	-0.180	0.074	0.912
equal prime:rise - equal prime:fall	-0.064	-0.206	0.078	0.873
female bias prime:rise - equal prime:fall	-0.039	-0.166	0.089	0.985
male bias prime:rise - equal prime:fall	-0.125	-0.259	0.010	0.093
male bias prime:fall - female bias prime:fall	0.068	-0.052	0.189	0.669
baseline:rise - female bias prime:fall	0.071	-0.041	0.183	0.537
equal prime:rise - female bias prime:fall	0.060	-0.068	0.188	0.849
female bias prime:rise - female bias prime:fall	0.085	-0.027	0.198	0.296
male bias prime:rise - female bias prime:fall	-0.001	-0.121	0.119	1.000
baseline:rise - male bias prime:fall	0.002	-0.117	0.122	1.000
equal prime:rise - male bias prime:fall	-0.008	-0.143	0.126	1.000
female bias prime:rise - male bias prime:fall	0.017	-0.103	0.137	1.000
male bias prime:rise - male bias prime:fall	-0.069	-0.196	0.058	0.719
equal prime:rise - baseline:rise	-0.011	-0.138	0.117	1.000
female bias prime:rise - baseline:rise	0.014	-0.097	0.126	1.000
male bias prime:rise - baseline:rise	-0.072	-0.191	0.048	0.607
female bias prime:rise - equal prime:rise	0.025	-0.103	0.153	0.999
male bias prime:rise - equal prime:rise	-0.061	-0.196	0.074	0.873
male bias prime:rise - female	-0.086	-0.206	0.034	0.368

Appendix. G. 2. Tukey HSD comparisons of average hit rates for rises between prime condition and speaker genders.

Comparison: Hit rates for rises	Difference	Lower CI	Upper CI	Adjusted p-value
equal prime:female speaker - baseline:female speaker	-0.023	-0.205	0.158	1.000
female bias prime:female speaker - baseline:female speaker	0.091	-0.067	0.248	0.658
male bias prime:female speaker - baseline:female speaker	-0.101	-0.270	0.069	0.614
baseline:male speaker - baseline:female speaker	0.013	-0.145	0.171	1.000
equal prime:male speaker - baseline:female speaker	0.014	-0.166	0.195	1.000
female bias prime:male speaker - baseline:female speaker	-0.051	-0.210	0.108	0.979
male bias prime:male speaker - baseline:female speaker	-0.030	-0.198	0.139	0.999
female bias prime:female speaker - equal prime:female speaker	0.114	-0.069	0.296	0.557
male bias prime:female speaker - equal prime:female speaker	-0.078	-0.271	0.115	0.924
baseline:male speaker - equal prime:female speaker	0.036	-0.146	0.219	0.999
equal prime:male speaker - equal prime:female speaker	0.037	-0.165	0.240	0.999
female bias prime:male speaker - equal prime:female speaker	-0.028	-0.211	0.156	1.000
male bias prime:male speaker - equal:female speaker	-0.007	-0.199	0.186	1.000
male bias prime:female seaker - female bias prime:female speaker	-0.192	-0.362	-0.021	0.015 *
baseline:male speaker - female bias prime:female speaker	-0.077	-0.237	0.082	0.821
equal prime:male speaker - female bias prime:female speaker	-0.077	-0.258	0.105	0.906
female bias prime:male speaker - female bias prime:female speaker	-0.141	-0.301	0.019	0.129
male bias prime:male speaker - female bias prime:female speaker	-0.120	-0.290	0.050	0.382
baseline:male speaker - male bias:female speaker	0.114	-0.057	0.285	0.462
equal prime:male - male bias prime:female speaker	0.115	-0.077	0.307	0.606
female bias prime:male speaker - male bias prime:female speaker	0.050	-0.121	0.222	0.987
male bias prime:male speaker - male bias prime:female speaker	0.071	-0.110	0.252	0.933
equal prime:male speaker - baseline:male speaker	0.001	-0.181	0.183	1.000
female bias prime:male speaker - baseline:male speaker	-0.064	-0.224	0.096	0.929
male bias prime:male speaker - baseline:male speaker	-0.043	-0.213	0.127	0.995
female bias prime:male speaker - equal prime:male speaker	-0.065	-0.247	0.118	0.961
male bias prime:male speaker - equal prime:male speaker	-0.044	-0.235	0.147	0.997
male bias prime:male speaker - female bias prime:male speaker	0.021	-0.150	0.192	1.000

Appendix. G. 3. Tukey HSD comparisons of average false alarm rates for rises between prime condition and utterance contour types.

Comparison	Difference	Lower CI	Upper CI	Adjusted p-value
equal prime:fall - baseline:fall	-0.017	-0.139	0.105	1.000
female bias prime:fall - baseline:fall	-0.125	-0.233	-0.018	0.010 *
male bias prime:fall - baseline:fall	-0.027	-0.142	0.087	0.996
baseline:rise - baseline:fall	-0.092	-0.199	0.016	0.160
equal prime:rise - baseline:fall	-0.131	-0.254	-0.008	0.027 *
female bias prime:rise - baseline:fall	-0.003	-0.111	0.105	1.000
male bias prime:rise - baseline:fall	-0.052	-0.167	0.063	0.871
female bias prime:fall - equal prime:fall	-0.109	-0.231	0.014	0.127
male bias prime:fall - equal prime:fall	-0.011	-0.139	0.118	1.000
baseline:rise - equal prime:fall	-0.075	-0.198	0.048	0.581
equal prime:rise - equal prime:fall	-0.114	-0.251	0.022	0.177
female bias prime:rise - equal prime:fall	0.014	-0.110	0.137	1.000
male bias prime:rise - equal prime:fall	-0.035	-0.165	0.094	0.991
male bias prime:fall - female bias prime:fall	0.098	-0.017	0.213	0.161
baseline:rise - female bias prime:fall	0.034	-0.075	0.142	0.982
equal prime:rise - female bias prime:fall	-0.006	-0.129	0.118	1.000
female bias prime:rise - female bias prime:fall	0.122	0.013	0.231	0.015 *
male bias prime:rise - female bias prime:fall	0.073	-0.042	0.189	0.537
baseline:rise - male bias prime:fall	-0.064	-0.179	0.050	0.687
equal prime:rise - male bias prime:fall	-0.104	-0.233	0.026	0.226
female bias prime:rise - male bias prime:fall	0.024	-0.091	0.140	0.998
male bias prime:rise - male bias prime:fall	-0.025	-0.147	0.097	0.999
equal prime:rise - baseline:rise	-0.039	-0.162	0.084	0.979
female bias prime:rise - baseline:rise	0.089	-0.020	0.197	0.207
male bias prime:rise - baseline:rise	0.040	-0.076	0.155	0.968
female bias prime:rise - equal prime:rise	0.128	0.004	0.252	0.037 *
male bias prime:rise - equal prime:rise	0.079	-0.051	0.209	0.591
male bias prime:rise - female	-0.049	-0.165	0.067	0.907

Appendix. G. 4. Tukey HSD comparisons of average false alarm rates for rises between prime condition and speaker genders.

Comparison	Difference	Lower CI	Upper CI	Adjusted p-value
equal prime:female speaker - baseline:female speaker	-0.020	-0.195	0.156	1.000
female bias prime:female speaker - baseline:female speaker	0.109	-0.045	0.264	0.386
male bias prime:female speaker - baseline:female speaker	0.066	-0.099	0.231	0.928
baseline:male speaker - baseline:female speaker	0.031	-0.123	0.184	0.999
equal prime:male speaker - baseline:female speaker	-0.028	-0.202	0.147	1.000
female bias prime:male speaker - baseline:female speaker	0.099	-0.055	0.253	0.515
male bias prime:male speaker - baseline:female speaker	0.045	-0.119	0.208	0.992
female bias prime:female speaker - equal prime:female speaker	0.129	-0.047	0.306	0.340
male bias prime:female speaker - equal prime:female speaker	0.086	-0.100	0.271	0.856
baseline:male speaker - equal prime:female speaker	0.051	-0.125	0.226	0.988
equal prime:male speaker - equal prime:female speaker	-0.008	-0.202	0.187	1.000
female bias prime:male speaker - equal prime:female speaker	0.119	-0.057	0.295	0.448
male bias prime:male speaker - equal:female speaker	0.064	-0.120	0.249	0.965
male bias prime:female seaker - female bias prime:female speaker	-0.043	-0.210	0.123	0.993
baseline:male speaker - female bias prime:female speaker	-0.079	-0.233	0.076	0.786
equal prime:male speaker - female bias prime:female speaker	-0.137	-0.313	0.039	0.259
female bias prime:male speaker - female bias prime:female speaker	-0.011	-0.166	0.145	1.000
male bias prime:male speaker - female bias prime:female speaker	-0.065	-0.230	0.100	0.934
baseline:male speaker - male bias:female speaker	-0.035	-0.200	0.130	0.998
equal prime:male - male bias prime:female speaker	-0.094	-0.278	0.091	0.787
female bias prime:male speaker - male bias prime:female speaker	0.033	-0.132	0.198	0.999
male bias prime:male speaker - male bias prime:female speaker	-0.021	-0.196	0.153	1.000
equal prime:male speaker - baseline:male speaker	-0.059	-0.233	0.116	0.972
female bias prime:male speaker - baseline:male speaker	0.068	-0.086	0.222	0.883
male bias prime:male speaker - baseline:male speaker	0.014	-0.150	0.177	1.000
female bias prime:male speaker - equal prime:male speaker	0.127	-0.048	0.302	0.355
male bias prime:male speaker - equal prime:male speaker	0.072	-0.111	0.256	0.934
male bias prime:male speaker - female bias prime:male speaker	-0.054	-0.218	0.110	0.974