Distribution Agreement

Signature:

______________________________ __________

Sabrina K. Sidaras Date

Hearing What You Expect to Hear:
The Interaction of Social and Cognitive Mechanisms
Underlying Vocal Accommodation

By

Sabrina K. Sidaras
Doctor of Philosophy

Psychology

Lynne C. Nygaard, Ph.D.
Advisor

Larry Barsalou, Ph.D.
Committee Member

Harry Gouzoules, Ph.D.
Committee Member

Laura Namy, Ph.D.
Committee Member

Donald Tuten, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the Graduate School

Date

Hearing What You Expect to Hear:
The Interaction of Social and Cognitive Mechanisms Underlying Vocal Accommodation

By

Sabrina K. Sidaras
Bachelor of Science, University at Buffalo, State University of New York, 2004
Master of Arts, Emory University, 2006

Adviser: Lynne C. Nygaard, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the Graduate School
of Emory University in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Psychology
2011

Abstract

Hearing What You Expect to Hear:
The Interaction of Social and Cognitive Mechanisms Underlying Vocal Accommodation

During spoken communication, individuals accommodate or change the way they speak based on characteristics of their conversational partner (Giles & Powesland, 1975; Giles, Scherer, & Taylor, 1979; Miller, 2005). Evidence suggests that accommodation may be due to a fundamental perceptual-production link (Fowler & Galantucci, 2005; Pardo, 2006; Pardo & Remez, 2006) that results in automatic vocal alignment with an interlocutor. However, social motivations have also been proposed as the primary underlying mechanism for accommodation (Giles, 1973; Giles & Coupland, 1991) because a variety of social factors have been shown to significantly influence the degree and type of vocal accommodation that occurs. This dissertation was designed to investigate how social variables might interact with underlying perception-production representations and mechanisms in speech vocal accommodation. The set of experiments sought to determine a) if adult listeners could perceptually identify characteristics of talker's voice such as speech rate and age that were used as indices to measure vocal accommodation and b) if perceptual-motor representations are accessed continuously and automatically during the perception of vocal stimuli even in situations when social context is unclear or ambiguous and finally c) whether social expectations about characteristics of a talker's voice, in this case age, can affect accommodation in a minimally social context. The findings from the first experiment showed that listeners could perceptually identify speech rate and age, and vocal accommodation occurred even in a minimal social context, but was not necessarily affected by implicit social variables. The second experiment investigated the extent to which social variables influence the perceptual-motor processing of speech when these variables were highlighted. Participants were primed with social stereotypes about age that have been shown to be reflected in speaking style and then were asked to shadow or repeat words produced by an age ambiguous speaker. Illusory accommodation did occur such that participants accommodated towards an expected vocal behavior rather than vocal characteristics actually present in the acoustic signal. These findings have implications for how social mechanisms interact with perceptual-motor representation and processing during vocal accommodation.

Hearing What You Expect to Hear:
The Interaction of Social and Cognitive Mechanisms Underlying Vocal Accommodation

By

Sabrina K. Sidaras
Bachelor of Science, University at Buffalo, State University of New York, 2004
Master of Arts, Emory University, 2006

Advisor: Lynne C. Nygaard, Ph.D.

A dissertation submitted to the Faculty of the Graduate School
of Emory University in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Psychology
2011

Acknowledgements

I would like to thank my advisor, Lynne Nygaard for her excellent mentorship throughout this process as well as throughout my graduate career. Her time, advice, knowledge, and support have been a driving force underlying my growth and development as a scientific researcher. She has been an extremely positive role model as a scientist and scholar, and has helped me think critically and deeply about many issues in our field. She was my support over the years both within and outside the lab. I am eternally grateful for all the time she has put in and her genuine efforts in helping me succeed. I couldn't have asked for a better mentor. The lessons she has ingrained with me will last far beyond my academic years and I will carry them with me and pass them on to others. I can only hope that this type of relationship between mentor and student exists among other institutions as its contributed to not only to my own positive experience, but also led to my gratitude towards Emory University as whole. Over the years, the confidence in my abilities has grown largely due to Lynne. Without her encouragement, advice, constructive criticism, and passion for the field, I would have not believed in my own abilities to become a responsible scientist or scholar.

I would like to thank my dissertation committee members, Laura Namy, Larry Barsalou, Harold Gouzoules, and Donald Tuten for their genuine enthusiasm and support from the moment I proposed and as I finished. Their time in this process is greatly appreciated as well as their feedback and questions. They have largely contributed to my interest in continuing this line of research. Laura has been my second advisor encouraging me along the way and spending a large amount of her time with feedback and suggestions on various projects I have worked on. She was not obligated in any way to be another advisor for me and I am sincerely and deeply appreciative to have had her guidance. Some students are extremely lucky to have one great advisor let alone a secondary advisor that is there for them as well. As for the Cognition and Development department as a whole, I cannot begin to express how fortunate I am to have such a collaborative and supportive faculty that not only got along academically but also interpersonally. It is often hard to describe how thankful I am to have a department that genuinely wants you to succeed and be happy doing it. I am fortunate for the department staff; Emily Stills and Paula Mitchell especially for their help in many avenues making sure the necessary processes have been completed to reach my degree over the years. The individuals previously mentioned are what made the journey of achieving such an honorable degree such a positive experience.

I am thankful for having a supportive colleague, Jessica Alexander-Duke for her encouragement. Her friendship helped pull me through many challenges. Zak Schroerlucke has been my best friend for over the past year reminding me of the passion and pride I should take both in my work and in my personal life. He has shown me that life is truly what you make of it. Many thanks are due to the researchers who helped in stimulus and data collection, especially Lauren Clepper, Megan Watters, Arielle Walzer, Lana Greenbaum, and Valentin Lazar as well as other students who have been contributors to the lab in the past. Their contribution is often overlooked but without their help, this project would not have been completed in the time that it did. And finally to my parents, who loved me the best way they knew how and kept my ego grounded and humbled always.

Table of Contents

I. Introduction ……………………………………………………………... 1

a. Evidence for Perception Production Links…………………………….. 4

b. Perception Production Link Account 15

c. Social Motivations Underlying Vocal Accommodation……………… 19

d. Social Expectations and Stereotypes……………………….…………24

e. The Current Investigation……………………………………………..29

II. Experiment 1a ………………………………………………………… 32

a. Method ………………………………………………………………... 33

i. Participants …………………………………………………….. 33

ii. Stimuli ……………………………………………………………. 34

iii. Procedure ……………………………………………………… 35

b. Results and Discussion ……………………………………………... 36

i. Overall Analyses…………………………………………….…….36

ii. Analyses of Individual Speakers…………………………………...39

III. Experiment 1b ……………………………………………………...... 43

a. Method ……………………………………………………………… 45

i. Participants …………………………………………………….. 45

ii. Stimuli ……………………………………………………………. 45

iii. Procedure ……………………………………………………… 45

b. Results and Discussion ……………………………………………… 47

i. Duration measurements……………………………………………..47

ii. Response latencies…………………………………………………48

IV. Experiment 2 ……………………………………………………………....... 50

a. Method ………………………………………………………….………52

i. Participants …………………………………………….…….. 52

ii. Stimuli… ……………………………………………………….. 53

iii. Procedure ……………………………………………………… 55

b.Results and Discussion ……………………………………………….56

i. Duration measurements……………………………………………56

ii. Response latencies………………………………………………..57

V. General Discussion ……………………………………………………...60

VII. References …………………………………………………………..79

VIII. Appendix …………………………………………………………..105

List of Tables and Figures………………………………………………………….

Table 1. *Average Durations and Standard Deviations of Stimulus Words Produced by each Old and Young Talkers at Fast and Slow Speaking Rates*

Table 2. *Means and Standard Deviations for Proportion Correct Perceptual Identification of AGE as a Function of Age of the Speaker and Speaking Rate*

Table 3. *Means and Standard Deviations for Proportion Correct Perceptual Identification of SPEAKING RATE as a Function of Age of the Speaker and Speaking Rate*

Figure 1. *Identification Performance for Age as a Function of Age of the Speaker and Speaking Rate (Error bars represent standard error).*

Figure 2. *Identification Performance for Speech Rate as a Function of Age of the Speaker and Speaking Rate (Error bars represent standard error).*

Figure 3. *Age Identification Performance for Individual Old and Young Speakers as a Function of Speaking Rate (Error bars represent standard error).*

Figure 4. *Speech Rate Identification Performance for Individual Old and Young Speakers as a Function of Speaking Rate (Error bars represent standard error).*

Figure 5. *Average Difference Scores across Age and Speech Rate (Error bars represent standard error).*

Figure 6. *Average Response Latencies across Age and Speech Rate (Error bars represent standard error)*

Figure 7. *Average Difference Scores for Accommodation Performance Following Old versus Young Primes (Error bars represent standard error).*

Figure 8. *Average Response Latencies for Accommodation Performance Following Old versus Young Primes (Error bars represent standard error)*

The way in which humans communicate through spoken language highly influences the quality of social interactions and the relationships created or maintained between conversational partners (interlocutors). Communication, as defined in the literature, occurs when signals carry information between a sender and a receiver (Krauss, 2002). These signals are highly complex, however, carrying information to the listener not only about *what* is being said, but *how* it is being said. How a speaker shapes his or her verbal utterance is just as important as the linguistic information being transmitted. Interlocutors often display behaviors in vocal communication that can function to improve the effectiveness, intelligibility, and ease of interaction as well as influence the social interaction and social relationships between partners (Berger & Bradac, 1982; Giles, Coupland, & Coupland, 1991; Giles & Powesland, 1975; Giles, Scherer, & Taylor, 1979, Triandis, 1960). One of the ways interlocutors can modify both the communicative and social effectiveness of the interaction with their conversation partners is through a behavior described as vocal accommodation.

Vocal accommodation refers to behaviors in which individual speakers adjust or change their vocal characteristics relative to their conversational partner. Vocal accommodation typically refers to changes in the indexical and nonverbal characteristics of spoken language. Indexical or talker-specific characteristics provide information about attributes of the talker such as individual identity, gender, emotional state, dialect, and socioeconomic status and are instantiated in acoustic properties like fundamental frequency (pitch) that may differ across different talker's voices or that may differ from utterance to utterance within a talker as well (Abercrombie, 1967; Frick, 1985; Labov, 1972; Van Lancker, Kreiman, & Emmorey, 1985). Nonverbal characteristics of communication can also include other properties of spoken language such as number of pauses, response latency or time between speaker switches, speech duration and speaking rate (Giles & Robinson, 1990; Scherer, 1979). Traditionally, indexical properties have been thought to be independent from those properties such as word use or grammatical structure that specifically involve linguistic structure in the language (Abercrombie, 1967; Joos, 1948).

Research has shown that talkers converge or become more like their conversational partner on a variety of vocal dimensions such as pronunciation (Giles, 1973; Pardo, 2006), speaking rate (Webb, 1970; Giles, Coupland, & Coupland, 1991), vocal pitch (Gregory, 1990, 1994; Gregory, Green, Carrothers, & Dagan, 2001; Gregory & Hoyt, 1982; Gregory & Webster, 1996; Gregory, Dagan, & Webster, 1997), pause and utterance duration (Mattarazzo, 1973) and intonation patterns (Gregory, 1990; Fernald, Taeschner, Dunn, & Papousek, 1989; Lieberman, 1967). More recently, Pardo (2006) has shown that talkers also accommodate to the fine structure of phonetic form such as the acoustic-phonetic properties of vowels produced by specific talkers. Although acoustic phonetic properties have not been explicitly defined as a property of indexical variation, acoustic phonetic form is still a component of talker variation that can be considered specific to an individual talker's voice.

Two primary classes of theories have been proposed to account for how and why vocal accommodation in spoken language communication occurs. Perceptual or cognitive accounts rest on the assumption that a link between perception and production may operate as a relatively automatic low-level process that allows accommodation to occur (Fowler, Brown, Sabadini, & Wihing, 2003; Fowler & Galantucci, 2005, Galantucci, Fowler, & Turvey, 2006; Shockley, Sabadini, & Fowler, 2004). This type of account suggests that the perception and production of speech are not independent from one another. Accommodation occurs even in nonsocial settings where social dialogue is not made explicit because of this low-level, presumably automatic perceptual-motor mechanism. In contrast, social accounts rest on the main principle that socially motivated constructs underlie vocal accommodative behavior so that accommodation most reliably occurs in explicit social situations (Coupland & Giles, 1988; Giles & Coupland, 1991; Shepard, Giles, & LePoire, 2001). Vocal accommodation can be considered part of a larger model in communication (i.e. Communication Accommodation Theory) that invokes a social affiliative account for why speakers will adapt to the communicative style of their conversation partner (Gallouis, Ogay, & Giles, 2005; Giles, 1971; 1973; Giles & Coupland, 1991; Miller, 2005). Traditionally, this type of account primarily assumes that the mechanisms underlying vocal accommodation are socially driven.

The main objective of this dissertation was to investigate how perceptual-motor and social constraints are integrated during vocal accommodation. By examining how social context and social affiliative motives interact with perceptual-motor processing and representation during accommodation, these experiments were designed to determine, 1) what type of circumstances must exist for social context to influence lower level perceptual and cognitive processes, 2) how automatic or explicit the processes may be that are involved in vocal accommodation, and 3) what kind of representations may be involved in the ability to execute this type of behavior. The results of this investigation were intended to contribute to our understanding of not only why and in what contexts vocal accommodation occurs, but also to clarify the role of both social and cognitive mechanisms in vocal accommodation.

Evidence for Perception-Production Links

Although not traditionally posited as a mechanism for vocal accommodation, recent research suggests that vocal accommodation may rely on low-level perceptual-motor links in behavior and as a consequence, be the result of the general processing architecture of the listener (Fowler, 2003; Fowler & Galantucci, 2005; Fowler, Galantucci, & Saltzman, 2003). This general fundamental perceptual motor link may consist of either common representational structures or common processing systems that result in a speaker's ability to mimic the perceptual details of the conversational partner's speech. This proposed link assumes that perception and production of speech are not independent from one another. Examples from both speech perception and production, and the cognitive and neuroimaging literature more generally, provide evidence for a link between perception and motor planning in the execution of human behavior and point to three possible ways that perception might be linked to production; 1) low level mappings between perception and production, 2) common representational domains between perception and production, and 3) motor simulations between perception and production. The evidence for each possible type of perception-production link is found for both speech specific behaviors and for other general motor behaviors not specific to speech.

Evidence for simple low-level mappings between perception and production have largely been examined in speech research with human adults (Fowler & Galantucci, 2005; Krauss & Pardo, 2006; Pardo & Remez, 2006). Research in second language learning has shown that changes in one domain (perception) can affect the other domain (production). When listeners are trained to perceive certain speech sounds in a second language, they also become better able to produce speech sounds in the second language

(Bradlow, Yamada, Reiko, Pisoni, & Tohkura, 1999; Bradlow, Pisoni, Yamada, Reiko, & Tohkura, 1997). For example, monolingual Japanese speakers have a difficult time differentiating between the English phonemes, /r/ and /l/, both perceptually and in production (Lively, Pisoni, Yamada, Tohkura, & Yamada, 1994; Logan, Lively & Pisoni, 1991). Bradlow et al. trained monolingual Japanese speakers to identify the /r/ and /l/ contrast in English. In addition to pre and post training measures of identification, Japanese participants' productions of /r/ and /l/ were recorded before and after training. By the end of training, listeners were not only better able to identify English productions of /r/ and /l/, but participants' own productions of /r/ and /l/ also improved significantly. The authors argued that the perceptual learning of the phonemes /r/ and /l/ transferred to the production of /r/ and /l/ phonemes. Improvement in the perception of certain speech sounds was related to improvement in the production of those speech sounds highlighting a possible link between perception and production. Similarly, Adank, Hagoort, and Bekkering (2010) found that when listeners were presented with an unfamiliar accent and asked to imitate the unfamiliar accent, their perception and comprehension of the accent improved relative to conditions when listeners were asked to simply repeat the unfamiliar accent without imitating, suggesting a benefit for imitating the accent rather than just engaging in the motor act of speaking. These findings show that producing the particular form of speech sounds can have an influence on perceiving speech sounds with similar properties. Finally, Ito and Ostry (2010) found that participants' perception of particular speech contrasts was affected when participants' facial skin was artificially stretched in a manner similar to stretching that occurs during movement of the lips and mouth to produce speech. Participants were presented with ambiguous speech sounds produced between the pronunciation of the words “head” and “had”. When participants’ facial skin was stretched upward consistent with the position of articulators when producing the word “head”, participants perceived the ambiguous speech sound to sound more like “head” than “had”. Likewise, when participant’s facial skin was stretched downward consistent with production of the word “had”, participants perceived the ambiguous speech sound to sound more like “had” than “head”. Taken together, these findings demonstrate that simple low-level mappings between perception and production of speech are tightly coupled so that changes in one domain have an effect on the other domain.

In addition to behavioral studies, findings from neuroimaging studies have shown that neural substrates for perception and production of speech sounds overlap, supporting the notion that perception and production may be linked by common representational domains (Hickok, 2001; Imada, Zhang, Cheour, Taulue, Ahonene, & Kuhl, 2006). For instance, many studies have shown activation in Broca’s area (located in the inferior frontal gyrus) during both speech production tasks and during speech perceptual tasks including passive listening to speech sounds and processing tasks involving phonological perception (Bookheimer, 2002; Binder et al., 2000; Zatorre, Evans, Meyer, & Gjeddi, 1996). Wilson, Saygin, Sereno, and Iacoboni (2004) were specifically interested in targeting motor areas in the brain to examine the view that parts of the motor system are activated in both the perception and production of speech. Wilson et al. asked participants to passively listen to monosyllables and then also to produce these same speech sounds. When participants listened to speech sounds, activation occurred in the superior portions of the ventral premotor cortex that largely overlapped with primary motor areas that were activated during the production of those same speech sounds. This finding suggests that the neural substrates for perception and production of speech share a common representational domain instantiated in the brain. Skipper, Nusbaum, & Small (2005) found that when participants both heard and saw a speaker (audiovisual perception) produce speech, activation occurred in brain areas associated with speech production suggesting that the same areas activated during speech production are activated during the perception of speech. In addition, others have found activation in the superior temporal lobes (part of the auditory cortex) during both the perception and production of indexical and non-indexical properties of vocalizations (Belin, Zatorre, Lafaille, Ahad & Pike, 2000; Fecteau, Armony, Joanette & Belin, 2004; Hickok, 2001; Hickok, Buchsbaum, Humphries, & Muftuler, 2003; Ozdemir, Norton, & Schlaug, 2006). Hickok et al. (2003) found that areas in the left superior temporal sulcus (STS) and the left and right premotor cortex were activated when participants were asked to listen and produce both speech and melodic tone sequences. When participants were asked to listen to tone sequences and then asked to produce them by humming, pronounced activation was found in the right portions of the STS. Findings from these neuroimaging studies suggest that the processing involved during the perception and production of vocalizations seem to rely on activation in overlapping areas of the brain providing evidence that perception and production may be linked through a common representational or processing format during the perception of speech.

Another way a fundamental link between perception and production may be instantiated is through motor simulations that are engaged both during perception and production of an action (Fowler, Brown, Sabadini, & Weihing 2003; Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Liberman & Mattingly, 1985; Wilson, Saygin, Sereno, & Iacoboni, 2004). For example, when an individual perceives speech, motor simulations in the form of articulatory gestures (physical motor movements) involved in the production of those same speech sounds are engaged to uncover the linguistic structure in the speech signal. This notion of motor reenactment during the perception of speech implies a type of simulation, which is defined as a process that activates previous actions that are stored in memory (Decety & Ingvar, 1990). Motor simulations involve activating motor areas of the brain that are involved during a normal motor action but doesn't cause overt movement (Hesslow, 2002). Thus, it is a simulation in the sense that an individual doesn't explicitly perform overt motor behavior but rather activates motor representations involved in producing that behavior.

Studies using Transcranial Magnetic Stimulation (TMS), another method of mapping brain activity, have revealed a common currency between perception and production that may reflect motor simulations (Fadiga, Craighero, Buccino, & Rizzolatti, 2002; Watkins, Strafella, & Paus, 2002; Watkins & Paus, 2003). TMS is a noninvasive procedure that involves placing an electromagnetic coil on the scalp and administering a magnetic pulse into the cortex to induce currents in the brain. Motor evoked potentials (MEP) are recorded from muscles following TMS of the motor cortex. Watkins and colleagues applied TMS to the face area of the primary motor cortex to elicit motor-evoked potentials in the lip muscles of participants. When participants listened to speech or viewed speech related lip movements, MEP's were larger than when participants listened to nonverbal sounds (e.g. glass breaking, guns firing, bells ringing) or viewed movements not related to speech (eye and brow movements). Fadiga et al., (2002) found similar increased MEP's of tongue muscles when listeners perceived speech sounds that implicated tongue movements versus speech sounds that did not involve tongue movements suggesting that when listeners were presented with speech sounds, they activated or simulated the same motor movements used in the production of those sounds. These findings demonstrate perceptual-motor links that underlie the perception and production of speech may take the form of motor simulations.

Another way in which perception might be linked to production for speech is through a simple priming mechanism. Priming is defined as an increased sensitivity for certain stimuli due to some prior experience involving the stimuli, which then facilitates a behavioral response (Jacoby, 1983). For instance, Pickering and Garrod (2004) have argued that during conversation interlocutors become aligned at phonological, syntactic and semantic levels. According to Pickering and Garrod, this alignment occurs as a result of a simple priming mechanism in which perception of linguistic levels in speech automatically primes representations associated with the production of language at each of the levels that are involved during conversation. Thus this alignment occurs as a result of priming between both the perception and production of speech during conversation between talkers.

So far, I have only reviewed evidence for links between perception and production in the domain of speech, however this capacity may be a general one that operates across motor and perceptual domains. For example, Prinz (1997) proposed that perception and production may be linked through a common representational domain for the perception and action of general motor behaviors. That is, when the activation of perception and motor codes occur simultaneously, the domains necessarily interact. Niedenthal, Brauer, Halberstadt, and Innes-Ker (2001) found that preventing a motor action that is consistent with what is being simultaneously perceived interferes with perception. Participants were shown a movie clip of faces that transitioned from sad to happy expressions and were asked to identify at what point in the clip a change occurred from the initial expression to a new expression. Half of the participants were told to perform the task with a pen in between their lips and teeth preventing any sort of facial imitation while the other half of participants performed the task without any obstruction. Participants who were able to imitate the faces identified the offset of the initial expression earlier than those participants who were prevented from executing a facial imitation (pen in between lips and teeth). This finding suggests that the production of motor behavior was involved in the perception of facial expression. Niedenthal et al. argued that preventing a motor response or the possibility of a response potentially interfered with perception because of the overlapping representation of the production and perception of the motor behavior.

Others have found that there is an overlap in processing both the perceptual representation of a motor behavior and the execution of the motor action. (Brass, Bekkering, & Prinz, 2001; Brass, Wohlslager, Bekkering, & Prinz, 2000; De Jong, 1993; Dejong & Sweet, 1994; Musseler & Hommel, 1997; Musseler & Prinz, 1996; Musseler, Wuhr, & Prinz, 2000; Pashler, 1994; Pazzaglia, Pizzamiglio, Aglioti, 2008; Prinz, 1997; Wilson & Knoblich, 2005). For example, Brass et al., (2000) found that participants are faster at producing a specific finger movement when they perceive the same finger movement than when they perceive an arbitrary finger movement that is not compatible with what is being produced. Similarly, others have shown that the ability to produce a motor action is influenced by the compatibility of the perception of the motor action (Kilner, Paulignan, & Blakemore, 2003; Drost, Rieger, Brass, Gunter, & Prinz, 2003; Grossjean, Zwickle, & Prinz, 2009; Zwickle, Grossjean, & Prinz, 2010) suggesting a commonality in processing between perception and production. Hamilton, Wolperth, and Frith (2004) found that participants' perception of an observed action was influenced by their production of that action. For example, participants were asked to judge the weight of a box lifted by an actor while actively lifting either a light or heavy box. Perceptions of the weight of the box was influenced by the action produced by the participant so that the weight of box was judged to be lighter when the participant actively lifted the lighter box and was judged to be heavier when the participant actively lifted the heavier box. Hamilton et al argued that these findings were a result of overlapping systems between action and perception. Behavioral studies have provided evidence for a coupled link between the perception and production of general motor behaviors that operate during simulation. Action can be primed by perceiving certain perceptual objects with action possibilities that are consistent with that particular action (Castiello, Lusher, Mar, Edwards, & Humphreys, 2002; Tucker & Ellis, 1998, 2001; Witt, Kemmerer, Linkenauger, & Culham, 2010; Zwaan, Stanfield, & Yaxley, 2002). In terms of motor simulation, Prinz (1997) suggests that when individuals perceive motor behaviors that share features with a similar action stored in their motor repertoire, they are more likely to produce or imitate that action suggesting that this type of priming mechanism is not only restricted to speech. For example, Castiello et al. (2002) asked participants to observe a task using a grasping action on an object. Participants were presented with a series of objects and were asked to grasp either the same object that had been observed or a different object. The object was either small or large in size. Observations that had compatible responses were labeled as prime trials whereas observations that had incompatible responses were unprimed trials. For example, a primed trial occurred when the participant observed grasping on a small object and then executed an action on a small object. Unprimed trials occurred when the participant observed grasping on a small object and then executed an action to the large object. Overall, different measures on the grasping action were faster for primed trials than unprimed trials. When participants were presented with an object that was similar in size to the object that was observed, their simulations of the actions on the observed object facilitated their motor response for the object that was presented to them. Participants' motor actions in the task were primed by the simulations of the observed action.

The notion that a link exists between perception and production for general motor behaviors is consistent with neuroimaging findings that have specifically targeted neural activity in motor areas of the brain (Buccino, Binkofski, Fink, Fadiga, & Fogassi, 2001; Hari et al., 1998; Pulvermüller, 2005, Rizzolatti, Fadiga, Matelli, & Bettinardi, 1996; Watkins, Strafella, & Paus, 2003; Watkins & Paus, 2004). For example, Buccino et al., found that when participants actually view bodily motor actions performed by another individual, activation occurs in the motor areas of the brain associated with these actions. Participants viewed mouth, foot and hand actions performed by another individual. Functional neuroimaging revealed activation in the premotor cortex and superior parietal areas, which are areas associated in the production of the observed actions. Participants appeared to simulate motor actions already built into their motor repertoire, and these simulations mapped onto areas of the brain associated with both the perception and production of motor actions related to different parts of the body. Pulvermüller (2005) found that simulations are formed when participants are visually presented with motion words. Participants were presented with motion words that were related to parts of the body (i.e. lick, pick, kick). When reading these types of motion words, activation was found in respective motor areas of the brain associated with the execution of those bodily actions suggesting that participants simulated the movement of body parts associated with the motion words. Thus, simulating an observed action either by perceiving the action itself or being presented with stimuli that are associated with the action activates areas in the brain also involved in the execution of that behavior. Taken together, these studies demonstrate that the processes involved in both the perception and production of behavior are instantiated in overlapping motor areas in the brain during the simulation of motor actions.

One way in which motor simulations either for speech or other types of perceptual-motor behaviors might be instantiated is through a mirror neuron like neural system. Research using single cell recordings, which identify individual firing of cells related to behavior, has revealed that a set of neurons referred to as mirror neurons in nonhuman primates become active during the production of goal-directed motor actions as well as during the perception of these same motor behaviors when performed by others (Fadiga, Fogassi, Povesi & Rizzolatti, 1995; Ferrari, Rizzolatti & Fogassi, 2005, Kohler et al., 2002; Rizzolatti, & Arbib, 1998; Rizzolatti & Craighero, 2004; Rizzolatti, Fogassi, & Gallese, 2001).

Presently there is no method for single cell recordings in humans. However, using various neuroimaging techniques, many researchers have found that the same brain areas are activated when individuals produce actions themselves, imitate an action, and observe these same actions performed by others (Chong, Cunnington, Williams, Kanwisher, & Mattingley, 2008; Dinstein, Hasson, Rubin, & Heeger, 2007; Kilner, Neal, Weiskopf, Friston, & Frith, 2009; Iacoboni, Woods, Brass, Bekkering, Mazziotta & Rizzolatti, 1999; Nishitani and Hari, 2000, 2002). Iacoboni et al. (1999) found that the same brain areas, Broca's area, right anterior parietal regions, and right parietal operculum, when participants were asked to produce finger movements they observed (imitation), when asked to produce the same finger movements in an unrelated task (execution of movements) and when asked to observe another individual making the finger movements. Even more intriguing, activation increased significantly when participants imitated the motor action than when they either observed or executed the motor action. Similar results were found by Nishitani and Hari (2000) using magnetoencephalography (MEG), which is an imaging technique that measures magnetic fields produced by the electrical activity in the brain. Participants were asked to grasp and manipulate an object (execution), observe the same actions performed by an experimenter (observation) or imitate the observed action (imitation). During observation and imitation, the pattern and sequence of activity in the frontal areas of the brain (Broca's area) were similar as for the actual production of the same movements. Nishitani and Hari (2002) found similar results for the imitation of lip movements so that the pattern and sequence of activity in Broca's area were similar for the production of lip movements and the observation and imitation of lip movements. These findings add to the body of evidence suggesting that common motor simulations are engaged both during perception and production of an action (Prinz, 1997).

Taken together, there is abundant evidence to suggest that a fundamental perceptual motor system underlies a variety of behaviors, those specific to speech and those general to other motor behaviors. A link between perception and production may exist through simple low-level mappings between perception and production, common representations or processing, and/or simulation of perceived motor behaviors. Evidence for how perception and production may be linked across motor behaviors could offer one possibility for how a general cognitive mechanism could underlie vocal accommodation. That is, perception production links that are not necessarily specific to speech, but are evident across a variety of general motor behaviors, could account for how vocal accommodation occurs.

Perception Production Link Account

A general perception production link provides a fundamental cognitive mechanism for vocal accommodation. In order for vocal accommodation to occur, there must be some matching of a perceived vocal behavior with a motor representation for that behavior to execute the same or similar vocal production. A fundamental connection between perceived speech characteristics and speech production would allow an individual to produce utterances that accommodate to the vocal behaviors of the conversation partner. This ability assumes that listeners must be able to attend to the properties of the talker's voice in order to make changes to their own verbal behavior. Research on the perception of spoken language demonstrates that listeners are not only perceptually sensitive to talker-specific characteristics of speech (Magnuson & Nusbaum, 2007; Mullennix, Pisoni, & Martin, 1989; Mullennix & Pisoni, 1990; Nygaard & Pisoni, 1998; Nygaard, Sommers, & Pisoni, 1993; Palmeri, Goldinger, & Pisoni, 1993; Pisoni, 1994; Pisoni & Martin, 1989; Nygaard, Sommers, & Pisoni, 1994; Remez, Fellowes, & Rubin, 1997; Yonan & Sommers, 2000), but can also detect changes in indexical characteristics that occur during vocal accommodation (Feldstein & Welokwitz, 1978; Giles & Powland, 1975; Matarazzo, 1973; Street, 1984; Webb, 1970). This perceptual ability is a processing prerequisite for individuals to change their vocal behavior based on the vocal behavior of their interlocutors.

A perceptual motor link may operate as an automatic unconscious process (Chartrand & Bargh, 1999) allowing accommodation to occur in situations where social context is ambiguous or unclear (Babel, 2009; Fowler, Brown, Sabadini, & Weihing, 2003; Goldinger, 1998; Namy, Nygaard, & Sauerteig, 2002; Pardo, 2006; Shockley, Sabadini, & Fowler, 2004). For example, Namy et al., asked participants in the laboratory to perform a word shadowing task in which participants repeated as quickly as possible a list of words presented over headphones produced by two male and two female talkers. As a consequence of the nature of the task, participants weren't engaging in any sort of social interaction with talkers but simply responding to the stimuli in the task. Before performing the shadowing task, participants were asked to produce the list of words presented to them visually on a computer screen and these utterances were used as baseline comparisons. A separate group of listeners were then asked to judge whether the baseline productions or shadowed productions sounded more like the original talker's utterances. Namy et al. found that participants' shadowed utterances were judged to be more similar than the participants' own baseline utterances to the original talkers' utterances.

Similarly, Shockley, et al. (2004) asked participants in the laboratory to shadow a list of words spoken by a model over headphones. Shockley, et al. manipulated the voice onset time (VOT) of the model's target words by lengthening the time between when the consonant was released and when voicing, the vibration of the vocal folds, began (Cho & Ladefoged, 1999). Participants were asked to produce the set of words before (baseline measure) the shadowing task so the participant's baseline productions could be compared to participant's shadowing productions. Acoustical analyses revealed that participants' shadowed words had longer VOTs than baseline words. Shockley, et al. argued that participants' VOTs converged to those of the manipulated target words. These findings suggest that some kind of low -level perceptual motor link operates automatically during accommodation so that the mere perception of an auditory stimulus maps onto a production of the appropriate and similar motor behavior without any explicit social interaction occurring.

Although there is evidence to support a basic perceptual motor link as an underlying mechanism for vocal accommodation, it remains unclear how social and situational factors might influence this presumably low-level automatic mechanism. Ultimately, vocal accommodation operates during conversations between interlocutors and is embedded in and supports social interaction. Conversation between interlocutors is the primary situation in which vocal accommodation occurs suggesting that social factors are an inherent part of the context or environment during interaction between talkers. Thus, how do social factors influence the possible perceptual-motor mechanisms that have been shown, at least in part, to underlie vocal accommodation behavior?

Previous research has demonstrated that vocal accommodation does not always result in an exact match to the vocal behavior produced by an interlocutor (Giles, Mulac, Bradac, & Johnson, 1987; Pardo & Remez, 2006), which may offer one piece of evidence that accommodation is not entirely a simple automatic low-level process. Partial accommodation can result in a change in vocal behavior that approaches, but does not match, the vocal behavior of the interlocutor. For example, a speaker may initially exhibit a speaking rate of 50 words per minute and change to match another speaker's rate of 100 words per minute. This change would constitute full accommodation. However, a speaker may instead change their rate to 75 words per minute, exhibiting partial rather than full accommodation (Street 1982). Krauss and Pardo (2006) also note that acoustic-phonetic output during accommodation is rarely an exact match to what is being perceived. The degree of accommodation that occurs may depend on several factors, although it has not been clearly assessed in the literature. Previous research has shown that accommodation across multiple dimensions of vocal behaviors can be perceived as being disinterested, patronizing or exaggerating (Williams & Giles, 1996; Coupland, Coupland, Giles & Henwood, 1988), which may be one hypothesis for why individuals choose to partially accommodate to their partner rather than fully accommodate. This would imply that individuals have control of the extent to which they accommodate suggesting that accommodation may not rely solely on an automatic perception production link. Others argue that accommodation goes beyond a match or imitation because it is a process that necessarily occurs through or in conjunction with an individual's own articulatory habits (Krauss & Pardo, 2004, Pardo, 2006). That is, true imitation can rarely occur since each talker has his or her own distinct individual talker characteristics that differ from any other talker. This account also limits a strict perception production mechanism suggesting that accommodation may not be so automatic and could be influenced by other intervening variables.

Pardo (2006) argued that individuals may use certain strategies to alter their own articulatory habits during vocal accommodation that go beyond simple priming, imitation, or simulation. Pardo asked pairs of participants to interact with each other and work on a map task that had verbal landmarks. One member's map included a path drawn from a starting point around various landmarks to a finishing point, and the companion's map contained only the starting point and landmarks. The goal of the task was for the pairs to work together so that the path on the first map, which could not be seen by the holder of the pathless map, could be duplicated on the second map. Completion of the task required participants to converse with one another and spoken samples of the landmark labels were collected both before, during, and after the conversational interaction. A separate set of listeners were asked to judge similarities in pronunciation between an individual's production of an item and their partner's production of an item. Convergence occurred when listeners judged a repeated item spoken by one talker in the task to be similar to the sample production spoken by the talkers' partner. Overall, the results showed that different degrees of accommodation were related to the gender of the individuals as well as the role they played in the task (either givers of information or receivers of information). Pardo concluded that the degree of accommodation was influenced by the social situation between interlocutors, the social roles of the interlocutors themselves, and possibly by the individuals' own articulatory constraints and habits, suggesting that phonetic convergence at least is a result of a dynamic process involving situational and other factors, rather than solely an automatic process driven by perception production links. This type of account may be a possible reason for why interlocutors display partial rather than full accommodation during social discourse and hints at a flexible process in which various factors influence accommodation behavior to a greater or lesser extent over time and across situations or contexts.

Social Motivations underlying Vocal Accommodation

An alternative to an exclusively low level perceptual motor account for vocal accommodation is a social affiliative account for accommodation. Communication Accommodation Theory (CAT) is perhaps the predominant social-affiliative account and was first introduced to explain why speakers shift their accents and code switch (alternate between two or more languages or dialects within a conversation) during spoken communication (Giles, 1973; Giles, 1984; Giles, Taylor, & Bourhis, 1973; Giles & Smith, 1979). More recently, CAT has been proposed as a general communication theory to explain why individuals adjust their nonverbal behaviors in general, and their vocal properties in particular to those of their interlocutors (Coupland & Giles, 1988; Giles & Coupland, 1991; Giles & Noels, 1997; Giles & Wadleigh, 1999). In general, CAT proposes that speakers adapt to the characteristics of their conversational partner in order to reach some type of social or relational goal through creating, maintaining, or decreasing social distance between themselves and their partners (Coupland & Giles, 1988; Bourhis & Giles, 1976, 1976; Giles & Coupland, 1991; Shepard et al., 2001).

One type of goal that is proposed to motivate vocal accommodation is the desire for social approval. Natale (1975) found that accommodation in vocal characteristics such as speaking rate, pause duration, response latencies, and loudness was related to individual differences in need for social approval. Participants that scored higher on a trait measure of need for social approval converged more to a partner's vocal properties than participants who scored lower on the need for social approval. A greater need for social approval can be driven by factors such as the probability of future interaction with the interlocutor, the degree of social power over the interlocutor or past exchanges/experiences with the interlocutor (Thakerar, Giles, & Cheshire, 1982). For instance, an individual may converge in order to maintain the social approval of an interlocutor that they will probably interact with again. The desire for approval is quite strong and can outweigh past interactions between interlocutors. For example, convergence could still occur if an individual's conversation partner did not converge in a past exchange but the individual still desires approval from their partner (Giles & Coupland, 1991). Research has shown that individuals seek social approval of those whom they rate as attractive (Byrne, 1971; Bourhis & Giles, 1976; Street & Giles, 1982). By converging on properties of vocal behavior, an individual can increase his or her own social attractiveness (Feldstein & Welkowitz, 1978; Putman & Street, 1984).

Power differentials between groups of people have also been shown to influence individuals' likelihood to accommodate. Speakers have been shown to accommodate to a greater extent to individuals perceived to have high prestige and power (Gregory, 1990, 1994; Gregory & Webster, 1996; Gregory, Dagan, & Webster, 1997; Lawson-Sako & Sachdev, 1996). For instance, Gregory and Webster asked participants to independently rate the social status of a number of different talk show guests who appeared on the Larry King Show. Talk show guests that were rated to have lower status than the host were found to converge more to the talk show host than those guests who were rated to have higher status.

Social status is the prestige an individual has within a group (Coats & Feldman, 1996) and this prestige is determined by the context or environment in which the interaction is taking place. The context thus defines who is more dominant in a group and who holds a high social status. The influence of social context has been found among talkers that often code-switch or alternate between two or more languages or dialects within a conversation (Lawson-Sako & Sachdiv, 1996). For example, in Tunisia, a country off the coast of North Africa, the majority of the people are bilingual. A very low percentage of people speak French as a first language, while the majority of people speak Tunisian

Arabic. Because French is rare as a first language in Tunisian culture, those that speak it are regarded to have a higher social status. For example, if the higher status bilingual speaker would converse with a lower status speaker, the lower status speaker would follow the language switches the higher status speaker produced.

Individuals may also vocally accommodate as a way to maintain or increase social rapport. Social rapport is often described as an affective state of social closeness that is characterized as a sympathetic or mutual understanding of trust or agreement between people. Previous work on vocal accommodation has shown that measures of social rapport are positively correlated with the degree of convergence that occurs between interlocutors (Feldstein & Welkowitz, 1978; Giles & Smith, 1979; Thakerar, Giles, & Cheshire, 1982) suggesting that individuals may attempt to align their vocal characteristics with those individuals they feel socially close to or want to feel close to. Feldstein & Welkowitz (1978) measured the desire for social closeness by asking conversational partners to rate each other on several dimensions of similarity. They found that the more similar an individual felt towards another person, the more socially close to their partner the individual felt. In turn, those who felt socially close with their conversational partner displayed more accommodation than those who did not rate themselves to be similar to their interlocutor.

Maintaining social identity during interaction is another motive that influences accommodation, but in the converse manner. Individuals will maintain or even amplify their own distinctive vocal behaviors to emphasize their own identity or highlight differences between their identity and that of their interlocutor (Cargile, Giles, & Clement, 1996; Kemp & Yaeger-Dror, 1991). Some researchers have argued that voice quality and speech style are among the most important markers of social identity (Giles, Scherer, & Taylor, 1979; Hummert, Mazloff, & Henry, 1999; Scherer & Giles, 1979). In order to maintain social identify, individuals will at times actually diverge away from the vocal behaviors of their conversation partner and emphasize their own vocal behaviors that are descriptive of their identity. Bourhis and Giles (1977) asked a group of participants who spoke both English and Welsh to participate in sessions surveying second language learning techniques. The experimenters screened participants to insure that they placed a strong value on their group membership and on the Welsh language (at the time only around 26% of Welsh could speak Welsh as opposed to English). The survey was presented to them in English by an experimenter who at one point during the session introduced a statement "Welsh was a dying language with no future" designed to threaten the participants' identity. Participants were found to increase their Welsh accents in their replies that followed the identity threat statement when compared to replies that were produced before the threat manipulation.

To summarize, the social affiliative account for accommodation assumes that talkers make adjustments to their vocal characteristics when interacting with other interlocutors due to several social motivating factors such as the desire for social approval, power differentials or social prestige among interlocutors, increasing social rapport, and maintaining social identity. All of these factors encompass some degree of social goal attainment such as increasing, decreasing, or maintaining social distance between interlocutors.

Social Expectations and Stereotypes

In regards to the research previously discussed, one body of work has examined social influences on vocal accommodation. Another body of work has addressed the underlying perceptual, motor, and cognitive processing structures that support this type of behavior. However, little or no work has addressed how these two accounts can be reconciled or how social variables might influence a basic perception production link during communicative interaction. Instead of either a cognitive or social account being an exclusive mechanism underlying vocal accommodation, it may be that these two types of mechanisms work together to influence the degree of accommodation that occurs.
Thus, rather than assuming that mechanisms for vocal accommodation rely primarily on lower level perceptual motor links or primarily on a dynamic social process, examining *how* social context and perceptual motor links *interact* may inform how both of these approaches account for the degree of accommodation that occurs. One way in which we could examine how social factors influence perceptual motor links underlying accommodation is through the extent to which social expectations or stereotypes have an effect on lower level perception production links. Although social expectations don't directly motivate a social goal per se such as the desire for social approval or maintaining social report, social expectations can influence how talkers make adjustments to their vocal characteristics when interacting with other interlocutors. Indeed, previous research has shown that listeners' accommodation behavior is influenced by social expectations about the talker (Fox & Giles, 1996; Gallois, 2004; Hummert, Garstka, Ryan, & Bonneson, 2004; Hummert, 1994; Montepare, Steinberg, & Rosenberg, 1992); however,

how these expectations influence lower level perceptual motor links has not been clearly addressed.

Social expectations are primarily expectations based on social norms and can be derived from stereotype formation. Social norms are shared ideas about what constitutes an appropriate or inappropriate behavior in any situation (Argyle, Furnham & Graham, 1981). For example, one type of social norm which is prevalent across a number of intergroup studies is that the minority group will converge toward a dominant outgroup (Gallois & Callan, 1991). Another social norm based on gender is that for the most part in many Western cultures, women accommodate more to their partners than men (Bilous & Krauss, 1988; Gallois & Callan, 1991).

One example of how social expectations or stereotypes might influence interactions with a communicative partner is the use of speech registers. For example, when younger interlocutors speak to elderly talkers, they often use “patronizing” speech or a register referred to as “Elder speech” that consists of a slower speech rate, simplified grammatical structure, and heightened articulation (Coupland, Coupland, Giles & Henwood, 1988; Giles, Coupland, Coupland, Williams, & Nussbaum, 1992; Kemper, 1994; McCann & Giles, 2006). Patronizing Elder Speech is often a result of expectations younger interlocutors have about elderly individuals being frail, slow, useless, inflexible or incompetent (Hummert, 1994; see Hummert, 1999 for a review; McCann & Giles, 2002; William & Giles, 1996). Because these characteristics do not apply to all older speakers, in some cases, younger individuals make false assumptions about the characteristics of older speakers, even when older speakers themselves may not actually demonstrate stereotypic characteristics. Furthermore, research has shown individuals will use this speech register even when there is no indication that the elderly person is suffering from a decrease in competency (Caporael & Culbertson, 1986; Montepare, Steinberg & Rosenberg, 1992). Thus, individuals adopt a particular speech register that is based on an expectation rather than sensitivity to the actual behavioral attributes of the older interlocutor. This use of speech register suggests that an individual can come into an interaction with a set of predispositions, beliefs or stereotypes about their communication partner and these stereotypes may in turn influence how an individual communicates with his or her interlocutor.

Research in domains other than intergenerational communication has demonstrated that listeners perceive characteristics of a speaker that are consistent with expectations they have about that speaker rather than the actual vocal behaviors that are produced. Rubin (1992) found that listeners who were presented with the same recorded passage perceived the speaker differently based on expectations. Participants, who were English-speaking psychology students, were asked to listen to a pre-recorded audio taped lecture produced by a native English speaker. One group was presented with a photograph of a Caucasian woman as they listened to the lecture whereas another other group was presented with a photograph of an Asian woman. When asked comprehension questions about the lecture, those students presented with the picture of the Asian woman made significantly more errors than those presented with the picture of the Caucasian woman. Thus, students who viewed the picture of the Asian person scored lower on comprehension of the lecture than those students presented with the picture of a Caucasian person. In addition, students rated the voice on the audiotape as having a stronger foreign accent when paired with the Asian photograph than when the voice was paired with the Caucasian photograph.

Thakerar and Giles (1981) provided information to establish the social status of a speaker to a group of listeners. The high status speaker was described as being highly competent based on a task administered in the study whereas the low status speaker was described as having low competency. The participants were given descriptions of the high and low status speaker but not explicitly told that the speaker was of low or high social status and then were presented with spoken passages. When the speaker was described as high status, participants perceived the speaker to speak faster with a more standard accent than they actually had (Thakerar & Giles, 1981). When the speaker was described as low status, participants perceived them to speak more slowly with a less standard accent than they actually had. In a later study, Thakerar, Giles, and Cheshire (1982) also found that interlocutors of different social status (both high and low) adopted a speech rate and changed their pronunciation to what they believed their partners should have, disregarding how their partner actually produced speech.

Taken together, these studies show that listeners appear to adopt a speaking style based on expectations, rather than automatically or reflexively alter their own vocal behavior towards what is actually being produced by the interlocutor. One possible explanation for how social variables might influence a simple perceptual motor link is that perception production links may operate during spoken language processing but social variables can influence the automatic operation of these perception production links. One circumstance that would lead social variables to influence automatic perceptual-motor processing is when expectations or stereotypes are accessed or simulated during a communicative interaction that then influence the type of accommodation that occurs. For instance, during the perception of a speaker's vocal behavior, several properties are stored in memory. These properties can include information about the social identity of the speaker as well as the acoustic properties of the talker's voice associated with that identity. Novel interlocutors that are encountered later and have some of these same perceptual properties and/or expected attributes can activate or prime social categories of identity which can include vocal behaviors associated with that identity. Stereotypes can then be accessed or simulated by the presence of physical features that are associated with the stereotyped group (Brewer, 1988; Pratto & Bargh, 1991). When these stereotypes are then simulated by the individual, the individual may adopt a vocal behavior that is associated with the stereotype and could accommodate based on this simulation rather than accommodate to the actual vocal characteristics being produced by the conversation partner. Thus, accommodation would be influenced by the stereotype rather than what is perceptually present in the acoustic signal.

In cases where individuals change their vocal behavior based on an expectation formed about their interlocutor (Thakerar & Giles, 1981), accommodation has typically been observed within a social interaction between interlocutors (but see Namy, Nygaard, & Sauerteig, 2002). Because communication is in itself a social interaction, social factors are inherently apparent in the situation. Thus, in the case where social expectations or stereotypes may have an influence on perceptual motor links during accommodation, it is difficult to assess whether the degree of accommodation that occurs is solely motivated by social goals or whether the degree of accommodation is a result of an interaction between social and lower level cognitive mechanisms. One way to examine these issues would be to minimize social context and observe whether or not individuals still accommodate. Thus, by minimizing social context, the notion that speakers are trying to achieve some social goal becomes less plausible and consequently, the degree of accommodation that occurs would not be due exclusively to social or motivational goals. Examining the influence of stereotypes on perceptual motor links in a task where social context is minimal may begin to identify how automatic perception production mechanisms are affected by social expectations during accommodation. If changes in vocal accommodation are due to some explicit motivation or intent on part of the interlocutor to reach some social goal, then effects of social variables should be minimal when participants are repeating words with little to no social context. Participants should accommodate based on the vocal cues that are present in the signal. However if changes in vocal accommodation are due to implicit social stereotypes, then accommodation due to social factors should persist even when there are no obvious social goals to be accomplished. If implicit stereotypes are induced, participants should accommodate based on expected vocal cues rather than on characteristics actually present in the physical acoustic speech signal demonstrating one way in which social factors can have an influence on automatic perception production links during accommodation. The goal of the present studies was to examine these possibilities.

The Current Investigation

The present study included two experiments that used age as a social factor in order to examine the influence of expectations on perceptual motor links during accommodation. Age was used because the age of a speaker has been found to be linked to speed of motor behavior in general, and to speaking rate in particular (Amerman & Parnell, 1990). In general, speaking rate has been identified as an age-related cue (Duchin & Mysak, 1987; Shipp, Huntley, & Hollien, 1992). Listeners appear able to estimate a speaker's age relatively well based on how the speaker sounds (Ptacek & Sander, 1966; Ryan &

Capadano, 1978) and older speakers are perceived to speak more slowly than younger speakers (Amerman & Parnell, 1992, Ryan & Burk, 1974). Amerman and Parnell asked listeners to assign pre-recorded passages produced by various speakers to a young, old, or dysarthric group (group in which speech is hard to understand due to being slurred, slowed, and having wide variation in tone usually a result of a neuromuscular disorder). In general, Amerman and Parnell found that listeners made significant errors by classifying normal elderly adults into the dysarthric group, perceiving normal elderly adults to speak much slower than younger adults suggesting that listeners have an expectation about the characteristics of talker's voice related to age. Speaking rate alone has also been associated with social expectations. Speakers that talk with a fast speech rate are judged to be higher in competence, sociability, and trustworthiness whereas speakers that talk with a slower speech rate are judged to be lower in competence (Scherer, 1979; Street & Brady, 1982, Street, Brady, & Putnam, 1983) suggesting that listeners associate speech rate with particular attributes of the talker. Another motivation for using age as a social factor is that others have found that age can be primed to simulate productions of behaviors associated with age stereotype (Bargh, Chen, & Burrows, 1996). Taken together, age seems to encompass robust stereotypes that could result in changes in motor behavior such as vocal accommodation. In addition, identification or perception of age has been linked with speech rate making it a useful index in measuring the degree of accommodation that may occur in relation to age of the talker.

The first experiment investigated if listeners could 1a) perceptually identify the age and speaking rate of various talkers and 1b) examined if perceived age has an effect on accommodation in a minimal social context. Experiment 1a was a perception task that examined whether listeners were able to perceive the actual characteristics in the speech signal and could identify the age of the talker as well as the speech rate of a talker. Separate groups of listeners made age identification judgments and speech rate judgments. In addition, Experiment 1a investigated whether age judgments are biased by speech rate or whether speech rate judgments are biased by age. For example, would listeners be more likely to identify fast items as young and slowly spoken items as old? Likewise, would listeners judge older talkers as always having a slow speaking rate than younger talkers even if speaking rate of older and younger talkers were similar? Experiment 1b was a production task to determine if perceived age has an effect on accommodation in a minimal social context. A naming task was used in which listeners were asked to repeat words spoken by older and younger speakers. Of interest was to observe whether listeners accommodate appropriately to older and younger speakers in a setting where social context was kept at a minimum. Would listeners accommodate to what is perceptually in the speech signal or would listeners' accommodation behavior be influenced by perceived age of the talker? Participants should accommodate based on the vocal cues that are present in the signal. However if changes in vocal accommodation are due to implicit social stereotypes that listeners form as a result of being able to identify age of the talker, then accommodation due to social factors should persist even when there are no obvious social goals to be accomplished.

Experiment 2 investigated whether priming of social expectations of age can influence accommodation to age-neutral speech. Participants who had been primed with young or old stereotypes were asked to repeat utterances produced by an age-ambiguous speaker. This type of ambiguity allowed us to examine whether participants accommodated to the actual age-neutral utterances or accommodated to an expectation of talker's voice with respect to primed age stereotype. Thus, even though all participants were presented with the same age-neutral utterances produced by an age-ambiguous speaker, half of the participants were primed with a young stereotype expecting the utterances to be produced by a young speaker, and half of the participants were primed with an old stereotype expecting the utterances to be produced by an old speaker. If stereotypes influence low-level perceptual motor links, then accommodation should be conditioned by social expectation, rather than solely by the characteristics that are actually perceptually present in the speech signal. This would suggest that accommodation in a minimally social setting isn't motivated exclusively from low-level perception production links but rather manifests from the interaction of social factors with perceptual motor links influencing the degree of accommodation that occurs.

Experiment 1a

The objectives of Experiment 1a were threefold. One objective was simply to confirm that listeners could identify the age of various talkers from single word utterances and could accurately judge speaking rate regardless of talker's age. If listeners are able to correctly identify the age of a talker and can identify the speech rate, these findings would demonstrate that listeners are perceptually sensitive to characteristics of talker's voice and can associate these characteristics with age. A second objective was to determine whether judgments of speaking rate and talker's age are independent. If listeners invoke stereotypes about older and younger adults when listening to speech, then speaking rate may influence judgments of age such that listeners would be more likely to classify someone as old if they are speaking slowly and less likely if they are speaking quickly. Likewise, age-related cues may influence judgments of speaking rate, such that listeners would be more likely to classify speaking rate as fast, if produced by a younger speaker and slow if produced by an older speaker. A third objective was to independently assess perceptual biases for the voices that were used to evaluate implicit social expectations or judgments on vocal accommodation in Experiment 1b. In addition to these objectives, Experiment 1a provided perceptual baselines for the same stimuli that were used in the naming task.

Experiment 1a employed a perceptual task in which listeners to words produced by both younger and older speakers at both fast and slow speaking rates. Half the listeners were asked to perceptually identify the age of the speaker (old vs. young) and half were asked to make a judgment on whether the item was produced at a fast or slow speaking rate.

Both younger and older speakers produced spoken items at a fast and slow speaking rates in order to determine whether listeners could correctly identify the age of the speaker, regardless of stereotypic characteristics in speaking rate associated with age. The objective of this experiment was to confirm whether or not listeners could identify age of the talker regardless of speaking rate and whether listeners could accurately identify speaking rate, regardless of the age of the speaker.

Method

Participants

Forty-eight Emory undergraduates (32 females and 16 males) were recruited from the Emory Research Pool to participate in this study. All participants were native speakers of American English with no history of a speech or hearing disorder and were between the ages of 19-25. Written informed consent was acquired from all participants under a research protocol that had been approved by the Emory University Institutional Review Board.

Stimuli

Four male talkers recorded stimuli for the experiment and their utterances were used in Experiment 1b as well. Only male talkers were used in the experiment in order to minimize any effect of the gender of the talkers on the degree of accommodation that might occur (see Namy, Nygaard, & Sauerteig, 2002). Two of the talkers were older in age (M=67.5 years) and two of the talkers were younger in age (M=24.5 years). Table I shows each individual talker's age. All talkers were Native English speakers.

The stimulus items were 40 bi-syllabic English words that were low frequency (M=37.12) based on Kucera & Francis (1967) norms. Low frequency words were used because previous research has shown that similarity between naming responses and baseline responses or imitation strength is stronger for low frequency words than for high frequency words (Goldinger, 1998). Each of the 4 talkers produced the 40 different words at a fast, medium, and slow speaking rate resulting in a total of 480 tokens. Only fast and slow speaking rate items were used for this experiment. Talkers recorded all stimuli at a medium speaking rate first and then produced all items at a fast and then a slow speaking rate. Talkers were encouraged to match their rates to a metronome set at 60 beats per minute to record neutral rates, 80 beats per minute for fast rates, and 40 beats per minute for slow rates The medium rate served as a baseline index for talkers during recording so that fast and slow items could be judged and produced relatively faster and slower than their neutral rate. The goal was to obtain utterances for which the average speaking rate of the slow items was comparable across both older and younger speakers and likewise, the average speaking rate of the fast items was comparable across both older and younger speakers.

Talkers' productions were recorded and digitized at 22.050 kHz sampling rate directly into a Dell computer using the Audacity audio recording program. Recorded utterances were then edited into separate files for presentation and amplitude normalized. Acoustic analyses were performed on each talker's productions to ensure that speaking rates were comparable, regardless of the age of the speaker. This was done by taking measures of average duration of items (in ms) produced by the older and younger speakers. Table 1 shows means and standard deviations of items produced by each of the older and younger speakers. There were no significant differences among speakers for average duration of fast items, $F(3, 117) = 0.58$, $p = 0.63$ *partial* $\eta^2 = .03$. Thus, fast items produced by each of the older speakers and each of the younger speakers were comparable to each other. Likewise, there were no significant differences among the speakers for average duration of slow items, $F(3,117) = 0.92$, $p = 0.44$, *partial* $\eta^2 = .07$. Thus, slow items produced by each of the older speakers and each of the younger speakers were comparable to each other.

Procedure

The experiment was run with E-Prime software (Schneider, Eschman, & Zuccolotto, 2002). Each participant was seated in a room in front of a computer. Participants were presented with all 40 stimulus items produced by the male speakers (2 older, 2 younger speakers) at fast and slow speaking rates. Half of the participants were asked to identify the age of the speaker, either 1) young or 2) old. The other half of the participants were asked to identify whether the item was produced in a 1) fast or 2) slow speaking rate. Each trial began with the presentation of a fixation cross on the screen for 1000 ms. Then, a spoken word was presented over headphones along with the two response options on the computer screen. Participants were asked to make responses using a button box and which option was presented on which side was counterbalanced across listeners. After a response had been made, there was a 500 ms interval between trials. Items were counterbalanced across speakers so that participants heard each speaker produce every item in a fast and slow rate resulting in a total of 320 items (40 fast items, 40 slow items produced by each of the 4 talkers). Items were only played once and were presented in random order. Participants were allowed to take optional breaks after every 80 trials.

Results and Discussion

Overall Analyses

To determine whether participants were generally sensitive to characteristics of talker's voice, one-sample t-tests were conducted to examine if listeners could correctly identify both age and speech rate at significantly above chance levels (M= 0.50). Table 2 shows mean proportion correct and standards deviations for judgments of age. Table 3 shows mean proportion correct and standards deviations for judgments of speaking rate. Participants were able to correctly identify the age of both older talkers, $t(23) = 6.63$, $p<0.001$, and younger talkers, $t(23) = 8.27$, $p<0.001$, significantly above chance. Similarly, participants were able to correctly identify speaking rate of fast, $t(23) = 15.28$, $p<0.001$, and slow items, $t(23) = 16.67$, $p<0.001$ significantly above chance. These findings show that participants were generally perceptually sensitive to surface characteristics of a talker's voice and could correctly identify attributes such as age and speech rate.

More interesting is the question of whether participants were biased by social expectations about age and speech rate when making each type of judgment. For example, if participants' identification of age of the talker was influenced by the speaking rate of the utterances, participants would be more likely to identify talkers as old when they were presented with slow items for both young or old talkers. Likewise, if participants' judgments of speaking rate were influenced by age, participants would be more likely to identify items as slow when presented with an older talker. In order to address this question, two separate 2x2 ANOVAs with Age (old vs. young) and Speech Rate (fast vs. slow) as within subject factors were conducted on age identification performance and speech rate identification performance. Analyses were conducted across participants (F_1) and items (F_2). Item analyses were conducted here and in subsequent analyses to confirm that the pattern of results observed across participants were also consistent across the items used in the study. For participants that made judgments about age, there was a main effect of age for participants and items, $F_1(1, 23) = 8.95$, $p < 0.01$, *partial* $\eta^2 = .28$, and F_2 $(1, 39) = 77.30$, $p<0.001$, *partial* $\eta^2 = .67$, a main effect of speech rate for participants and items, $F_1(1, 23) = 41.94$, $p < 0.001$, *partial* $\eta^2 = .65$, and F_2 $(1, 39) = 89.71$, $p<0.001$, *partial* $\eta^2 = .70$, and an interaction between age and speech rate for participants and items, $F_1(1, 23) = 47.54$, $p < 0.001$, *partial* $\eta^2 = .67$, and F_2 $(1, 39) = 666.56$, $p<0.001$, *partial* $\eta^2 = .95$,. Figure 1 shows age identification performance as a function of age of the speaker and the speaking rate of the tokens. Follow up analyses revealed that when participants made judgments about age, they were better at identifying the age of old speakers when items were presented at a slow rate (M=0.85, SD=0.12) than when presented at a fast rate (M=0.49, SD=0.19), $t_1(23) = 9.44$, $p<0.001$, and $t_2(39) = 20.79$, $p<0.001$. In addition, participants were better at identifying young speakers when the item was presented at a fast rate (M=0.88, SD=0.13), than when it was presented at a slow rate (M=0.72, SD=0.26), $t_1(23) = 3.61$, $p<0.001$, and $t_2(39) = 13.53$, $p<0.001$. Thus, when identifying age, participants seemed to be influenced by speech rate such that participants were better at identifying old speakers when items were produced at a slow rate than at a fast rate and better at identifying young speakers when items were produced at a fast rate versus a slow rate.

For participants that made judgments about speech rate, there was a main effect of age across participants and items, $F_1(1, 23) = 9.46$ $p<0.01$, *partial* $\eta^2 = .29$, and $F_2(1, 39) = 4.26$, $p<0.05$, *partial* $\eta^2 = .10$, a main effect of speech rate across participants and items, $F_1(1, 23) = 21.69$ $p<0.001$, *partial* $\eta^2 = .48$, and $F_2(1, 39) = 26.48$, $p<0.001$, *partial* $\eta^2 = .40$, and an interaction between age and speech rate across participants and items, $F_1(1, 23) = 14.70$ $p<0.01$, *partial* $\eta^2 = .39$, and $F_2(1, 39) = 9.17$ $p<0.01$, *partial* $\eta^2 = .19$. Figure 2 shows speech rate identification performance as a function of age of the speaker and the speech rate of the tokens. Follow up analyses revealed that when participants made judgments about speech rate, they were more accurate identifying fast items when the speaker was young (M=0.88, SD=0.12) than when the speaker was old (M=0.85, SD=0.13), $t_1(23) = -3.60$, $p<0.01$, and $t_2(39) = -3.80$, $p<0.001$. However, participants were not significantly more accurate identifying slow items when the speaker was old (M=0.91, SD=0.12) than when the speaker was young (M=0.91, SD=0.12),

$t_1(23) = 2.10$, $p = 0.87$, and $t_2(39) = 0.23$, $p = 0.82$. Thus when identifying speech rate, participants seemed to be influenced by age only for the fast items. Taken together, these patterns of results suggest that age judgments were more biased by speech rate than speech rate judgments biased by age.

These results suggest that listeners can identify attributes of the speech signal such as age of the speaker and speech rate. However, speech rate appears to be a salient perceptual property that listeners are sensitive to when making perceptual judgments about factors such as age of a speaker. Age does not seem to influence judgments of speaking rate to the same extent. Because speech rate seems to significantly influence listeners' ability to identify old and young age, it may act as a primary cue for perceivers to the age of a speaker. Interestingly, perhaps because speech rate is so salient, perceptual judgments about speech rate were not as influenced by other properties in the speech signal related to age.

Analyses of Individual Speakers

In order to examine the extent to which these effects were consistent across individual speakers, one sample t-tests were conducted to determine if listeners could correctly identify both age and speech rate significantly above chance (M=0.50) for each speaker. Table 2 shows means and standard deviations of identification performance for each speaker when making judgments about age. When making judgments about age, listeners correctly identified age above chance for speakers Young Male 1 (YM1), $t(23)= 7.74$, $p<0.001$, Young Male 2 (YM2), $t(23) = 6.60$, $p<0.001$, and Old Male 1 (OM1), $t(23) = 12.24$, $p<0.001$, but did not correctly identify age significantly above chance for Old Male 2 (OM2), $t(23) = 1.04$, $p=0.31$. Table 3 shows means and standard deviations of identification performance for each speaker when making judgments about speaking rate. When making judgments about speech rate, listeners correctly identified speech rate above chance for each speaker, Old Male 1, $t(23) = 15.94$, $p<0.001$, Old Male 2, $t(23) = 15.07$, $p<0.001$, Young Male 1, $t(23) = 15.79$, $p<0.001$, and Young Male 2, $t(23) = 17.39$, $p<0.001$. Listeners performed significantly above chance when making judgments about speech rate across each individual speaker. However, listeners did not perform significantly above chance when making judgments about age for one of the old male speakers.

In order to examine whether identification of age was more or less influenced by speaking rate across individual speakers, a 2x4 ANOVA with Speech Rate (fast and slow) and Speaker (OM1, OM2, YM1, YM2) as within subject variables was conducted on Age identification performance and Speech Rate identification performance. Analyses were conducted across participants (F_1) and across items (F_2). Figure 3 depicts age identification performance for each old (Old Male 1= OM1, Old Male 2= OM2) and young speaker (Young Male 1=YM1, Young Male 2=YM2) as a function of speaking rate. For participants that made perceptual judgments about Age, there was an effect of Speech Rate for participants and items, $F_1(1, 23) = 8.94$, $p<0.001$, *partial* $\eta^2 = .28$, and F_2 $(1, 39) = 77.30$, $p<0.001$, *partial* $\eta^2 = .67$, an effect of Speaker for participants and items, $F_1(3, 69) = 37.92$, $p<0.001$, *partial* $\eta^2 = .62$, and F_2 $(3, 117) = 87.17$, $p<0.001$, *partial* $\eta^2 = .69$, and an interaction between Speaker and Speech Rate for participants and items, $F_1(3, 69) = 34.08$, $p<0.001$, *partial* $\eta^2 = .60$, and F_2 $(3, 117) = 263.58$, $p<0.001$, *partial* $\eta^2 = .87$. Follow up analyses revealed that participants had more difficulty identifying age of speaker OM2's fast items (M=0.31, SD=0.20) than speaker OM1's fast items (M=0.68, SD=0.22), $t_1(23) = -10.17$, $p<0.001$, and $t_2(39) = -11.05$, $p<0.001$, suggesting that for fast items, it was easier to identify speaker OM1 than OM2 as old. Likewise, participants had more difficulty identifying age of speaker OM2's slow items (M=0.77, SD=0.22) than speaker OM1's slow items (M=0.94, SD=0.22), $t_1(23) = -4.31$, $p<0.001$ and $t_2(39) = -10.69$, $p<0.001$. There was no significant difference for identifying age of speaker YM1's fast items (M=0.86, SD=0.15) and speaker YM2's fast items (M=0.89, SD=.015), $t_1(23) = -0.85$, $p = 0.40$, and $t_2(39) = -1.85$, $p<0.07$. For slow items, there was a trend so that participants had more difficulty identifying age of speaker YM2's slow items (M=0.67, SD=0.30), than speaker YM1's slow items (M=0.78, SD=0.29) $t_1(23) = 1.87$, $p = 0.07$, and $t_2(39) = 5.08$, $p<0.001$, although this trend was not significant across participants. Taken together, these results suggest that identification performance of age especially for OM2 was significantly poorer than identification performance of age for OM1 when items were produced in a fast rate or slow rate. In general, when identifying age, the difference between fast and slow items was significantly different for each speaker, but the magnitude and direction of that difference changed as a function of individual speaker.

Figure 4 depicts rate identification performance for each old and young speaker as a function of speaking rate. For participants that made perceptual judgments about Speech Rate, there was a main effect of Speech Rate for participants and items, $F_1(1, 23) = 6.58$, $p<0.01$, *partial* $\eta^2 = .22$, and $F_2(1, 39) = 4.26$, $p<0.05$, *partial* $\eta^2 = .10$, a main effect of Speaker for participant and items, $F_1(3, 69) = 12.45$, $p<0.001$, *partial* $\eta^2 = .35$, and $F_2(3, 117) = 16.53$, $p<0.001$, *partial* $\eta^2 = .30$, and an interaction between Speaker and Speech Rate for participant and items, $F_1(3, 69) = 22.75$, $p<0.001$, *partial* $\eta^2 = .50$, and $F_2(3, 117) = 3.06$ $p<0.05$, *partial* $\eta^2 = .07$. Follow up analyses revealed that for speaker OM1, participants performed significantly better at identifying speech rate for slow items (M1=0.91, SD1=0.12) than fast items (M1=0.83, SD1=0.13), $t_1(23) = -5.23$, $p<0.001$, and $t_2(39) = -4.38$, $p<0.001$. For speaker OM2, participants performed significantly better at identifying slow items (M=0.92, SD=0.13) than fast items (M=0.89, SD=0.12), $t_1(23) = -2.36$, $p<0.05$, and $t_2(39) = -3.98$, $p<0.001$. For speaker YM1, participants performed significantly better at identifying speech rate for slow items (M=0.92, SD=0.11) than fast items (M=0.86, SD=0.12), $t_1(23) = -4.23$, $p<0.001$, and $t_2(39) = -4.88$, $p<0.001$. For speaker YM2, however, the pattern was reversed. Participants performed significantly better at identifying speech rate for fast items (M=0.92, SD=0.11) than slow items (M=0.90, SD=0.12), $t_1(23) = 2.17$, $p<0.05$, and $t_2(39) = -2.14$, $p<0.05$. Taken together, participants were better at identifying speech rate of slow items than fast items for speaker OM1, OM2, and YM1. For speaker YM2, participants were better at identifying speech rate of fast items than slow items. In general, when identifying speech rate, the difference between fast and slow items was significantly different for each speaker, but the direction of that change varied as a function of individual speaker.

These findings show that although listeners could correctly identify both Age and Speaker above chance *across* speakers, listeners had difficulty identifying the age of one speaker in particular, OM2, when making judgments about age in comparison to the other individual speakers in the study. These findings show that overall, listeners' age judgments were biased by speech rate. Listeners' judgments of age were influenced by how fast or slow the talkers produced items. However, judgments of speech rate were not as susceptible to variations in age. Speech rate seemed to be the primary driving factor for the perceptual identification of fast and slow items. In addition, the results suggest that although speech rate was a highly salient cue to talker age across individual speakers, speakers did vary in how well their age was identified. In particular, Old Male 2 speaker seemed to have ambiguous age characteristics that were particularly influenced by speaking rate. Listeners seemed to have trouble identifying the perceptual talker characteristics in order to make correct judgments about age for this speaker.

Experiment 1b

The objective of Experiment 1b was to examine whether vocal accommodation would occur in the absence of an explicit social context and to evaluate the extent to which readily perceived characteristics of a talker's voice, (i.e. age and actual speaking rate) could separately and together impact the degree of vocal accommodation in a low social context task. That is, would listeners systematically alter their rate of speech to approximate that of the presented speech or would differences in perceived age influence accommodation in a task that did not emphasize communicative interaction? Experiment 1b employed a production task in which listeners were asked to repeat or shadow words spoken by the same younger and older speakers used in Experiment 1a. The results from Experiment 1a showed that listeners could perceptually identify age of a speaker when listening to speech, albeit identification was significantly influenced by rate. Thus, to the extent that listeners are able to identify the age of the talker, then perceived age could influence accommodation to words produced at different speaking rates by older and younger talkers. If a listener uses information about perceived age of a talker and identifies the talker as an older speaker, this perception could influence how the listener accommodates to different speech rates produced by that older speaker. For example, if a listener is sensitive to the vocal characteristics of the older speaker, that listener might accommodate to the speaker based on a social stereotype that older individuals are less competent (speaking slower with a wider variation in tone from baseline) even though the older speaker does not display stereotypic vocal characteristics. This type of accommodation would suggest that social variables can influence the operation and automaticity of a perception production link underlying vocal accommodation. Listeners may not accommodate to what is perceptually present in the speech signal, but rather intervening social variables may influence the perception production link underlying accommodation behavior

Experiment 1b examined if age of the talker would influence talkers' accommodation to differences in speaking rate. Experiment 1b employed a naming or shadowing task in which listeners were asked to repeat words presented through headphones spoken by both younger and older speakers. Although others have used similar tasks, (Goldinger 1998; Namy, Nygaard, & Sauerteig, 2002), the proposed study evaluated the hypothesis that low-level perception-motor processing would operate during the perception and production of vocal stimuli in this task and lead to vocal accommodation. More importantly, this study evaluated whether speaker-specific expectations based on a social factor such as age would influence the extent to which listeners would accommodate in a non-social setting.

Methods

Participants

Forty-one Emory undergraduates (23 females and 18 males) were recruited from the Emory Research Pool to participate in this study. All participants were native speakers of American English with no history of a speech or hearing disorder and were between the ages of 19-25. Written informed consent was acquired from all participants under a research protocol that had been approved by the Emory University Institutional Review Board.

Stimuli

The same four male talkers from Experiment 1a were used for this experiment. As in Experiment 1a, two of the talkers were older in age and two were younger in age. The stimulus items consisted of the same 40 bi-syllabic English words that were used in Experiment 1a. Again, only fast and slow items were used in this study so that each of the four talkers produced 40 different items at a fast and slow rate. As in Experiment 1a, speech rates were comparable across old and young speakers.

In each condition, participants were presented with 40 shadowing trials in which each of the talkers produced five different fast items and five different slow items. Talker-word pairings were counterbalanced across participants and rotated through conditions so that every item appeared once with every talker across conditions.

Procedure

The design of the experiment consisted of two phases; a baseline task that used to measure the duration of participants' baseline utterances and a shadowing task to measure the duration of participants' shadowed utterances. Each participant was seated alone in a room in front of a computer and the experiment was run with E-Prime software (Schneider, Eschman, & Zuccolotto, 2002). In order to measure the participant's baseline utterances, participants were asked to read each of the 40 words that comprised the stimulus set. Each word was presented visually on a computer screen. Before each word was presented, a fixation cross was presented on the screen for 1000 ms and participants were asked to make their responses into microphones that were connected to the computer and used to record verbal responses. After participants made their response, there was a 500 ms interval between each trial. These items served as a baseline in order to measure each participant's average speaking rate. Individuals differ in their characteristic speaking rate (Allen & Miller, 1999; Miller, Grosjean, & Lomanto, 1984) and as such, these individual differences must be accounted for when examining the degree of accommodation that occurs relative to an individual's average speaking rate.

Before participants performed the baseline task, they were presented with five practice trials (English words not used in the rest of the experiment) in order to get them accustomed to making verbal responses into the microphones. After the baseline task, participants were asked to shadow or repeat isolated words spoken by each of the four talkers and presented over headphones. Participants were simply asked to repeat the word. On each trial, a fixation cross was presented on the screen for 1000 ms followed by the presentation of a spoken word. After the participant made a response into the microphone, there was a 500 ms interval between trials. Words were only played once. Similar to the baseline task, participants were presented with ten practice trials (English words not used in the rest of the experiment) before they began the shadowing task.

Participants were given no instructions to explicitly accommodate or produce the item in a similar way to that of the talker. Items were presented in random order.

Results and Discussion

Duration measurements. In order to assess the degree of accommodation, individual word durations were measured to determine the speaking rate of each shadowed utterance produced by each participant. All duration measurements were conducted using Praat sound analysis software (Boursma, Paul, & Weenink, 2008). Word durations were measured for both the participants' baseline and shadowed trials and baseline item durations were subtracted from the shadowed item durations of each speaker in order to control for variation in the length of particular items and for individual variation in characteristic speaking rate of the participant. A positive difference score indicates that a participant's shadowed response was slower in duration than their baseline response. In turn, a negative difference score indicates that a participant's shadowed response was faster in duration than their baseline response. In order to evaluate whether or not vocal accommodation occurred, these difference scores were compared across both old and young talkers producing both fast and slow speech rates. Because shadowing durations did not reliably vary as a function of individual talker, analyses on the difference scores across participants for each individual speaker was not examined further.

Figure 5 shows mean difference scores for old and young talkers as a function of speaking rate. In order to determine whether participants changed their vocal characteristics as a function of age or speaking rate, difference scores were analyzed using a 2 x 2 ANOVA with Age (old and young) and Speaking rate (slow and fast) as within subject variables. Analyses were conducted across participants (F_1) and across items (F_2). Item analyses were conducted to confirm that the pattern of results observed across participants were also consistent across the items used in the study. The analysis revealed there was a significant main effect of speech rate for participants and items, $F_1(1, 40) = 33.92$, $p<0.001$, *partial* $\eta^2 = .46$, and $F_2(1, 39) = 247.90$ $p<0.001$, *partial* $\eta^2 = .86$. Participants' shadowed utterances were longer in duration when repeating items that were produced in a slow rate relative to their baseline durations (mean difference score = 37.12) and were shorter in duration when repeating items produced in a fast rate relative to their baseline utterance durations (mean difference score = -31.88). However, no significant main effect of age for participants and items, $F_1(1, 40) = 2.27$ $p = 0.14$, *partial* $\eta^2 = .05$, and $F_2(1, 39) = 1.03$, $p = 0.32$, *partial* $\eta^2 = .32$, and no significant interaction between age and speech rate for participants and items, $F_1(1, 40) = 0.68$, $p = 0.41$, *partial* $\eta^2 = .02$, and $F_2(1, 39) = 0.73$, $p = 0.40$, *partial* $\eta^2 = .02$ were found suggesting that degree of accommodation did not change as a function of age.

Response latencies. Participants' response latencies to initiate the shadowing responses were also examined to determine if the pattern of results observed for participants' accommodation performance was consistent with the amount of time it took for the participant to initiate their shadowing responses. Thus, for example, when durations of participants' shadowed utterances were slower when repeating items that were produced in a slow rate relative to items produced in a fast rate, were participants' response latencies also slower in response to items that were produced in a slow rate? Response latency was measured from the beginning of each word to the moment the participant initiated a shadowing response.

Figure 6 shows response latencies across Age and Speech Rate. A 2 x 2 ANOVA with Age (old and young) and Speaking rate (slow and fast) as within subject variables was conducted on response latencies across participants (F_1) and items (F_2). The analyses revealed there was a significant main effect of speech rate for participants and items, $F_1(1, 40) = 242.97$, $p<0.001$, *partial* $\eta^2 = .86$, and $F_2(1, 39) = 257.19$, $p<0.001$, *partial* $\eta^2 = .87$ which was consistent with the results found for participants' duration measures. Participants took a longer amount of time to respond to items that were produced in a slow rate (M=965.09, SD=184.85) and took a shorter amount of time to respond to items produced in a fast rate (M=1292.92, SD=144.29). However, no significant main effect of age for participants and items, $F_1(1, 40) = 0.10$, $p = 0.75$, *partial* $\eta^2 = .003$, and $F_2(1, 39) = .002$, $p = 0.96$, *partial* $\eta^2 = .000$, and no significant interaction between age and speech rate for participants and items, $F_1(1, 40) = 0.07$, $p = 0.79$, *partial* $\eta^2 = .002$, and $F_2(1, 39) = 0.13$, $p = 0.72$, *partial* $\eta^2 = .003$ were found suggesting that participants' response latencies did not change as a function of age. These results were consistent with the findings from participants' duration measures. Thus, not only did participants' accommodation behavior change as a function of speaking rate, but the time it took for participants to initiate a shadowing response also changed as a function of speaking rate.

Taken together, participants demonstrated vocal accommodation regardless of talker's age. These findings reveal that accommodation did occur in a low social context suggesting that a perception production link is a relatively automatic continuous process that occurs doing the perception of vocal stimuli and may be driving vocal accommodation. However, perceived age of the speaker did not appear to influence the extent to which participants accommodated and as such, perceived age did not appear to have an intervening effect on perceptual motor links underlying accommodation. The degree of accommodation that occurred relied primarily on vocal cues such as speech rate that may be obvious to the listener. Speech rate was such a salient cue and such a salient cue to age that it may have masked or overwhelmed any the effects that speaker-specific expectations may have had. Still at issue is whether invoking social expectations and highlighting social variables may affect the degree of accommodation when other indexical characteristics are less salient or ambiguous, particularly in contexts in which social communicative interaction is kept at a minimum.

Experiment 2

Previous research has shown that it is possible to invoke stereotypes implicitly and that this type of stereotype can have later effects on behavior (Bargh, 1994, 1999; Brewer, 1988; Chartrand, Maddux, & Lakin, 2005; Chen & Bargh, 1997; Devine, 1989; Ferguson & Bargh, 2004; Smith, 1990; Patterson, 2001). For example, Dijksterhuis and van Knippenberg (1998) asked participants to imagine the behaviors, lifestyle, and appearance of either a professor or secretary for five minutes. Participants were then given a set of general knowledge questions. Those primed with the stereotype of college professor performed significantly better on the general knowledge test than those primed with the stereotype of secretary. In addition, those primed with the stereotype of secretary performed faster at completing the questionnaire than those participants primed with college professor and control participants who were not primed. In order to explore how social variables could influence a perception production link underlying vocal accommodation, an invoked stereotype about age was used to examine the influence of age stereotype on vocal accommodation.

We used this age stereotype in order to examine if participants' behavior in a minimally social task, in this case a shadowing task, would be affected by an expectation of talker's voice related to an age stereotype. As mentioned earlier, a body of research suggests that younger individuals possess stereotypes about the way older individuals speak. Younger and middle-aged individuals seem to be consistent in their stereotypes about older individuals being inflexible, slow, cognitively impaired, and timid (Hummert, 1994; Hummert, Shaner, & Garstka, 1998; Stewart & Ryan, 1982). In addition, some research has shown that listeners will associate how older speakers sound when producing speech with negative stereotypes about the speaker (Mulac & Giles, 1996). In addition, because younger individuals have preconceived expectations that older individuals are cognitively impaired, they adopt a particular speaking style with older individuals by speaking slower, inserting more pauses, and using more simple speech forms (Coupland, Coupland, Giles, & Henwood, 1988; Hewstone & Giles, 1986; Ryan, Giles, Bartolucci, & Henwood; 1986; Thakerar, Giles & Cheshire, 1982).

In the present experiment, invoked stereotype was a between subjects variable. An "old" or "young" age stereotype was primed to create an expectation about a talker's voice (older speaker - slower speaking rate, younger speaker - faster speaking rate). Participants were primed with social stereotypes about age, in this case "old age" or "young age" using fictional descriptions. Participants were told that the items they were about to hear in the shadowing task were produced by either the old or young talker they had been primed with. However, the actual utterances produced by talker in the experiment were neutral in speaking rate and neutral in judged age. The objective of this experiment was to investigate how expectations about a type of speaker and/or stereotypes about the speaking style of an individual would influence vocal accommodation of talker's utterances even when the utterances that were actually produced by talkers were age neutral.

By priming expectations about a talker's speaking style, it was possible to evaluate whether these social expectations would influence the automaticity of perceptual motor links to accommodate to stereotypical vocal characteristics that are not present in the acoustic speech signal. If illusory accommodation does not occur, and participants accommodate to what is perceptually present in the speech signal, this would provide evidence that accommodation in a minimal social task is motivated by lower level perception production links. However, if illusory accommodation does occur, and an invoked stereotype has an effect on participant's vocal behavior, it would demonstrate one way in which social factors could influence perceptual motor processes underlying accommodation. In addition, these findings could provide evidence that accommodation in a minimal social setting is not completely dependent on perceptual motor links and that accommodation occurs as a result of an interaction between social factors and cognitive perception production mechanisms operating relative to an invoked expectation about talker's voice.

Methods

Participants

Eighty-one Emory undergraduates (54 females and 27 males) were recruited from the Emory Research Pool to participate in this study. All participants were native speakers of American English and were between the ages of 19-25. Written informed consent was acquired from all participants under a research protocol that had been approved by the Emory University Institutional Review Board.

Stimuli

The goal of the present experiment was to present listeners with age and rate neutral utterances in order to evaluate the effects on accommodation of expectations about speaker characteristics. In Experiment 1a, listeners had difficulty identifying the age of one of the older speakers, Old Male 2, and as such this speaker was identified as a candidate speaker for the current experiment. In order to confirm that this speaker was judged to be of neutral age when producing neutral rate utterances, a pilot study was conducted in which the neutral rate utterances of speaker Old Male 2 as well as speaker Old Male 1, and two other speakers ranging in age from 40-45 years were presented to a group of 12 listeners. The mean for neutral utterances across speakers was 530.38 ms. Speaker Old Male 1 who had already been correctly identified as old, was used in order to assess if the ambiguity of speaker Old Male 2 was due to either items being produced in a fast or slow speech rate or to talker characteristics specific to that of Old Male 2. Additional talkers between the ages of 40-45 were used to determine if this age range would influence listeners' perceptual judgments of age. The additional talkers' productions were recorded as before, directly into a Dell computer digitized at a 22.050 kHz sampling rate using the Audacity audio recorder program. Utterances were edited into separate files for presentation and amplitude normalized. During pilot testing, participants were asked to identify age (old vs. young) for each of these speakers producing the same stimuli used in Experiment 1a and 1b in a neutral speech rate. As in Experiment 1a, listeners still had difficulty correctly identifying age for speaker Old Male

2 (M=0.44, SD=0.27), t(11)=-0.75, *p*=0.47. Likewise, as in Experiment 1a, listeners correctly identified age for speaker Old Male 1 significantly above chance (M=0.74, SD=0.19), *t*(11) = 3.64, *p*<0.01. For both of the speakers that was between the ages of 40-45, listeners identified age of the speaker to be old significantly above chance (Speaker 1, M=0.88, SD=0.28), *t*(11) = 4.72, *p*<0.01; Speaker 2, (M=0.72, SD=0.24), *t*(11) = 3.14, *p*<0.01. Thus, Old Male 2 was used as the ambiguous age-neutral speaker for this study since items produced by this speaker were neither identified as being produced by an old or young speaker. In addition, only items produced by Old Male 2 at a neutral speech rate were used for this study (M=543.05 ms).

The stimulus items consisted of the same 40 bi-syllabic English words that were used in Experiment 1a and Experiment 1b. The single talker's productions of all 40 items at a neutral speaking rate resulted in a total of 40 items heard by the listener.

Age primes

Pictures: A set of two pictures of males varying in age, one male that was 77 years of age and one male that was 22 years of age, was used to implicitly prime participants with age appropriate stereotypes (see Appendix B). These pictures were taken from a lifespan database of facial stimuli of neutral expressions produced by Minear & Park (2004).

Narratives: A set of two narratives of males varying in age, one male that was 70 years of age and one male that was 22 years of age were used to implicitly prime participants with age appropriate stereotypes. Each narrative was matched on number of words (word count=293) and described stereotypic behaviors that were age appropriate (see Appendix C). Thus, one narrative described behaviors associated with a young male and one narrative described behaviors associated with an older male. Each narrative was associated with a specific male name; “Tommy” for the young narrative and “Mr Jones” for the old narrative. In addition, each narrative informed the reader that these were individuals who had participated in previous research studies and had recorded auditory stimuli in the lab before. Examples of adjectives that were stereotypical of old were “cautious”, “careful”, “wrinkly”, and “grey”. Examples of adjectives that were stereotypical of young were “independent”, “energetic”, and “outgoing”. In general there was an attempt to make the narratives for old and young as affectively neutral as possible.

Procedure

The same procedures were employed as in Experiment 1b using a baseline and naming or shadowing task however in the naming task, participants were presented with items produced in a neutral speech rate by an ambiguous talker and were introduced to the speakers after baseline and before the experimental shadowing trials. Before participants performed the baseline task, they were presented with five practice trials (English words not used in the rest of the experiment) in order to get them accustomed to making verbal responses into the microphones. Between the baseline and naming task was a priming phase in which each participant was introduced to *either* a young or old stereotype by presenting the appropriate narrative and picture of “Tommy” or “Mr. Jones”. Thus old and young prime conditions were manipulated between subjects. Participants were asked to read the narrative quietly to themselves and then they were presented with the shadowing task. Similar to the procedures in Experiment 1a, participants were presented with ten practice trials (English words not used in the rest of the experiment). Before they began the shadowing task, participants were told that the words they were about to hear were produced either by “Tommy” or “Mr. Jones”. Participants were asked to simply repeat the word with no instruction to mimic or imitate the voice.

Results and Discussion

Duration measurements: All duration measures were conducted using Praat sound analysis software (Boursma, Paul, & Weenink, 2008) and carried out on each of the items produced by the participant. As such, word durations were measured for both the participants’ baseline and shadowed trials in order to compute difference scores to measure the degree of accommodation that occurred. As before, baseline item durations were subtracted from the shadowed item durations of each speaker in order to control for variation in the length of particular items and for individual variation in participants’ characteristic speaking rate.   Recall that a positive difference score indicates that a participant’s shadowed response was slower in duration than the participant’s baseline response. In turn, a negative difference score indicates that a participant’s shadowed response was faster in duration than the participant’s baseline response.

Figure 7 shows average difference scores across old prime and young prime conditions. In order to evaluate whether participants were accommodating differently when primed with old versus young, difference scores for participants primed with old versus young were compared. If the stereotype or expectations manipulation primed or activated biased perceptual representations, listeners should have accommodated to the appropriate stereotypic vocal behavior rather than to what was actually perceptually present (neutral rate items). If the participant was primed with an “old” stereotype and accommodated to the stereotypical behaviors of “old”, in general participants should have produced shadowed utterances that were longer in duration relative to their baseline utterances resulting in a positive difference score. If the participant was primed with a "young" stereotype and accommodated to the stereotypical behaviors of "young", in general participants should have produced utterances that were shorter in duration (e.g., faster speaking rate) relative to their baseline resulting in a negative difference score.

In order to assess whether the manipulation used in the present study influenced accommodation behavior, an independent sample t-test (t_1) for participants and for items (t_2) was conducted on difference scores for participants that were primed with "old" versus "young". The t-tests revealed that difference scores differed significantly across conditions for participants and items, $t_1(79) = 5.39$, $p<0.001$, and $t_2(78) = 14.59$, $p<0.001$. When participants were primed with "old", their shadowing responses slowed down relative to baseline (M=23.13) and when participants were primed with "young", (M=-28.09) their shadowing responses sped up relative to baseline (M=-28.09), These findings demonstrate that participants' speaking rates were slower when primed with an "old" stereotype than when primed with a "young" stereotype and faster when primed with a "young" stereotype than when primed with an "old" stereotype.

Response latencies. Participants' response latencies were analyzed to examine whether the pattern of results observed across participant's duration measures for stereotypes of "old" and "young", were also consistent with the amount of time it took for the participant to initiate their shadowing responses. Were participants' response latencies similarly affected by the primed stereotype? If so, then the time to initiate a shadowing response would be faster when participants are primed with a young stereotype than an old stereotype, and slower when participants are primed with an old stereotype than a young stereotype. If not, then the time to initiate a shadowing response should not differ significantly when participants are primed with a young stereotype or an old stereotype. If response latencies are consistent with the pattern of results found for participants' duration measures when primed with stereotypes of "old" or "young", it would suggest that the vocal accommodation in this case may be a general motor effect not specific to vocal accommodation. If response latencies do not differ as a function of the primed stereotype of old or young, it would suggest that priming may be specific to vocal accommodation and not a product of a general motor effect.

Figure 8 shows response latencies across old prime and young prime conditions. Response latency was measured from the beginning of each word to the moment the participant initiated a shadowing response. An independent sample t-test (t_1) for participants and for items (t_2) was conducted on the response latencies for participants that were primed with "old" versus "young". Interestingly, the t-tests revealed that response latencies did not differ significantly across conditions for participants and items, $t_1(79) = -0.51$, $p = 0.61$, and $t_2(78) = -1.06$, $p = 0.29$. Thus when participants were primed with "old", the time it took to initiate a shadowing response did not differ significantly than when primed with "young". The time it took for participants to initiate a shadowing response was not influenced by the stereotype manipulation even though participants' speaking rates during accommodation were slower in duration when primed with an "old" stereotype than when primed with a "young" stereotype and faster when primed with a "young" stereotype than when primed with an "old" stereotype. That response latencies did not vary as a function of the prime but accommodation behavior did strongly suggests that this invoked stereotype was not operating on motor behavior in general, but rather was operating to influence *speech* perception-production links specifically. These findings may also suggest that initial processing of the acoustic speech signal, which was neutral in rate and ambiguous in age, was not influenced by the social expectation of talker's voice because response latencies were not affected, but that social factors did have an influence on accommodation later during speech production.

Taken together, these findings suggest that participants activated a stereotype of "old" or "young" and accommodated to a speaking rate stereotypic of an "old" versus "young" talker regardless of the talker's actual utterances, which in the present experiment were produced at a neutral rate by the same ambiguous age talker. Participants appear to accommodate toward an expectation of talkers' voice, rather than to what was actually produced by the talker.

Notice that even though social variables were shown to influence a presumably relatively automatic perceptual-motor process, a fundamental perception production link still may be operating during accommodation. However, the present findings do suggest that this link is susceptible to other factors. Because these social variables were deliberately invoked and influenced accommodation in a setting with minimal social context where communication interaction was not taking place per se, these findings show that the instantiation of a perception production link does not override other factors such as an expectation about characteristics of a talker's voice based on an age stereotype. These findings also suggest that the degree of accommodation that occurred wasn't a direct result of exclusive social mechanisms since the shadowing task itself was not a socially interactive task where clear social goals could be identified by the participant. By attempting to minimize social interaction in the task, there were no obvious social goals to be attained and we observed that accommodation not only still occurred, but also that the degree of accommodation was influenced by a social expectation of talker's voice rather than what was perceptually presented to the listener. That is, participants showed illusory accommodation towards an expected vocal behavior rather than accommodation towards characteristics of talker's voice that were actually perceptually present in the acoustic speech signal. These findings suggest that social factors had an influence on perceptual motor processing during accommodation, which resulted in a type of accommodation that was illusory.

General Discussion

The experiments reported here investigated the social and cognitive mechanisms that underlie vocal accommodation. By examining the interaction of both social and perceptual factors during spoken language processing, it was possible to assess the degree of influence that social context may have on the lower level perceptual motor processes that appear to drive the adjustment of vocal properties during accommodation. To that end, the set of experiments assessed listeners' sensitivity to vocal characteristics of age and speaking rate as well as the extent to which these factors would influence vocal accommodation in a minimal social context. Across experiments, the results showed that adult listeners could perceptually identify speech rate and age, and vocal accommodation occurred even when listeners were simply asked to repeat words presented over headphones. In the first experiment, accommodation was not necessarily affected by implicit social variables. However, in the second experience, when participants were primed with social stereotypes about age and then were asked to shadow or repeat words produced by an age ambiguous speaker, illusory accommodation occurred. Participants accommodated towards an expected vocal behavior consistent with a social stereotype or expectation rather than to vocal characteristics actually present in the acoustic speech signal. These findings suggest that language users are both sensitive to the social significance of particular vocal characteristics and that when primed, expectations about vocal style can influence perceptual motor links during accommodation.

In order to assess the influence of social variables on perception production mechanisms, we had to first determine if listeners could perceptually identify surface characteristics of talker's voice that were used as indices for accommodation. In Experiment 1a, listeners were asked both to identify the age and speaking rate of utterances produced by old and young talkers, at fast and slow speaking rates. In general, listeners were able to correctly identify both types of characteristics of talker's voice. These results are consistent with previous findings showing that listeners are sensitive to nonlinguistic or surface characteristics of talker's voice and can identify and process this perceptually detailed information from the acoustic speech signal (Allen & Miller, 2004; Bradlow, Nygaard, & Pisoni, 1999; Nygaard, Burt, & Queen, 2000; Palmeri, Goldinger, & Pisoni, 1993; Van Lancker, Kreiman, & Emmorey, 1985.). However, the results also showed that age and rate did not appear to be processed independently. Listener's judgments of age were influenced by speech rate and likewise, judgments of speech rate were to some degree influenced by age.

That listener's judgments of age depended on how fast or slow the items were produced suggests that speech rate served as a particularly salient cue to age. This finding is consistent with previous research showing that younger listeners expect that older individuals will talk with a slower speech rate than younger individuals (Amerman

& Parnell, 1992; Coupland, Coupland, Giles & Henwood, 1988; Hummert, 1994), which may be an explanation for why the current listeners, who were younger in age (Emory undergraduates between the ages of 18-25) used speech rate when making judgments about age. Indeed, previous research has shown that listeners perceive older speakers to articulate and speak more slowly as they age (Ryan & Burk, 1974, Shipp, Qi, Huntley, & Hollien, 1992). These perceptions and expectations may be grounded in reality as others have shown that speech rate slows as individuals age. Older individuals speak more slowly than younger individuals both during spontaneous speech and when reading (Hartman & Danhauer, 1976; Ramig, 1983; Ryan, 1972).

Identification of speech rate was also affected by speaker's age. Listeners were significantly better at identifying rate for fast items spoken by young than by old talkers. Despite this influence of age on rate judgments, judgments of speech rate appeared to be more independent of the age of the talker than age judgments were of rate, suggesting that speech rate was in general a more salient property of the spoken utterances. One reason that speech rate may have been so salient is because only fast and slow speech rates were presented. Although both fast and slow speech rates were comparable across both young and old speakers, there was a relatively large difference between the average fast rate for talkers (M=393.55 ms) and the average slow rate for talkers (M=889.00 ms) highlighting the obvious differentiation between fast and slow speech rates to listeners.

Another reason may be that differences in the voice characteristics of old and young speakers may not have been as identifiable outside of rate, making age judgments more susceptible to variation along the other dimension. For example, listeners had difficulty correctly identifying age for the Old Male 2 (OM2) speaker leading this speaker to be ambiguous in perceived age. Although the other speakers' fast and slow rates were comparable to the fast and slow rates of this speaker, speech rate alone seemed to be the primary and most important factor in identifying age for OM2. Listeners appeared to be almost completely dependent on speaking rate, correctly identifying the age of OM2 only when presented with slow utterances. These results suggest that the difficulty in identifying age of OM2 could be due to ambiguity in characteristics of this talker's voice, other than speech rate, that index the age of a speaker. OM2 may have shared some characteristics of younger talkers' voices or had ambiguous properties that thus made it difficult to identify OM2 as either old or young. For the other speakers, however, speech rate seemed to bias but not entirely determine judgments of age. Many researchers have shown that there are several other properties inherent in talker's voice that differentiate old and young voices such as variability of pitch, breathiness and tremor, variability of voice onset time, jitter, shimmer, and spectral noise (Gorham-Rowan & Laures-Gore, 2006; Ptacek & Sander, 1966; Ramig & Ringel, 1983; Sweeting & Baken, 1982). In addition, age identification from talker's voice appears to be highly complex and can vary based on several factors like physiological condition (good or poor), and different types of speech samples such as spontaneous speech or reading (Schotz, 2003, 2007). Taken together, although speech rate is a characteristic that appears to be particularly indicative of age, it is not the only characteristic of talkers' voices that listeners are sensitive to or use as a cue to make age judgments. Thus, despite variability across talkers, the present study demonstrated that listeners can correctly identify both age and speech rate across multiple talkers.

In Experiment 1b, when listeners were asked to shadow or repeat spoken items by old and young talkers in fast and slow speech rates, they mimicked the vocal characteristics of each speaker's voice. These results are consistent with previous findings showing that accommodation can occur in a minimal social context outside of a conversational interaction (Goldinger, 1998; Goldinger & Azuma, 2004; Namy, Nygaard, & Sauerteig, 2002; Pardo, 2006; Shockley, Sabadini, & Fowler, 2004). Participants increased the speaking rate of their shadowed responses when repeating fast speaking rate utterances and decreased the rate of their utterances when repeating slow utterances, even though there was little interactive or communicative context, suggesting that accommodation was not solely due to social mechanisms. Perceptual-motor representations appeared to be accessed continuously and automatically during accommodation, which in turn allowed listeners to adjust their vocal characteristics of the presented words.

The results also showed that at least in a task in which social context or communicative interaction was minimized, there was little evidence that the social variable of age had an influence on the automatic perceptual motor links that appear to drive convergent vocal behavior observed during accommodation. Even though listeners could identify perceptual information about the age of the talker, this sensitivity to age did not have an effect on accommodation. Listeners sped up or slowed down their naming responses based exclusively on speech rate, with little influence of perceived age. Thus, listener's degree of accommodation was a direct result of what was presented to them in the acoustic signal. Any expectations based on the age of the speaker did not have an overall effect on the characteristics of their shadowing responses. These results suggest that by making social context minimal and social goals ambiguous, accommodation is largely driven by the automaticity of perception production links in speech.

One possible explanation for these findings is that because speech rate was such a salient property of these utterances and a clear cue to age as demonstrated in Experiment 1a, speech rate may have overridden any social influence of the age of the talker on accommodation. A listener's perceptual sensitivity to particular vocal properties in the acoustic speech signal obviously will constrain what properties listeners will accommodate towards. For example, from research in communication accommodation, listeners have been shown to both converge on some linguistic variables and diverge on others (Bilous & Krauss, 1988). Others have suggested that individuals will accommodate to linguistic features that are highly salient and obvious from the signal before accommodating to features that are not as salient (Trudgill, 1986). Thus, listeners may have accommodated to speech rate over perceived age of the talker primarily because it was such a salient perceptual cue.

Findings from participants' response latencies were consistent with the idea that speech rate was particularly salient to listeners. Participants' time to *initiate* a shadowing response was related to how fast or slow the item was produced by the talker, rather than to the perceived age of the talker. When participants heard an utterance produced in a fast rate, they initiated a shadowing response more quickly than when they heard an utterance produced in a slow rate, regardless of whether the utterance was produced by an old or young talker. Likewise, when participants heard an utterance produced in a slow rate, they initiated a shadowing response more slowly than when they heard an utterance produced in a fast rate, regardless of whether the utterance was produced by an old or young talker.

Taken together, these findings seem to suggest that listeners were responding to the information inherent in the actual acoustic speech signal, in this case, speech rate information, and mirror the findings from the perception experiment showing that perceived age was more biased by changes in speech rate than perceived speech rate was biased by changes in age. Even though participants were able to identify the age of the talker in Experiment 1b, their perception of age was influenced by speech rate information, reinforcing the notion that speech rate was a highly salient perceptual cue. Because speech rates were comparable across talker's age and listeners were not informed to explicitly think about age as a factor, their accommodation appeared to be based exclusively on the speech rates inherent in the acoustic signal. In addition, because the shadowing task itself is a relatively online processing task, listeners automatically accommodated to what they perceived. This, in turn, could have contributed to the automatic operation of perception production links without any influence of social variables resulting in accommodation to the actual acoustic signal.

Social factors did influence accommodation when a stereotype or expectation about age was invoked in Experiment 2. When participants were told to expect that the words they were about to hear were produced by an old or young talker, the results showed that the degree of accommodation was related to a formed expectation about the vocal behaviors of old and young talkers rather than to what was actually present in the acoustic speech signal. Thus, illusory accommodation occurred so that not only did listeners that were primed with an old stereotype slow down their shadowing responses, but also listeners that were primed with a young stereotype sped up their shadowing responses suggesting that the degree of accommodation that occurred as a function of primed expectations about age was symmetrical. Participants appeared to accommodate based on their expectations of young and old speech, even though all the utterances were neutral in speech rate and the task itself still involved no communicative interaction or clear social context. When listeners were encouraged to attend to a social expectation about the talker, social variables interacted with the automaticity of lower level perceptual motor links in speech to produce accommodation that was based on a social expectation rather than what was perceptually present in the signal.

One possibility for how social factors influenced accommodation behavior is that listeners may have simulated or embodied a general motor behavior related to age expectation that in turn influenced vocal accommodation. A listener's embodiment of an age stereotype may have led to the production of a general motor adjustment associated with age, which included vocal accommodation. Bargh, Chen, and Burrows (1996) found that when priming participants with stereotypes or trait constructs, participants enacted behaviors that were consistent with the content of the primed stereotype or construct in unrelated tasks. For example, in one of their experiments, Bargh et al. presented participants with words related to or unrelated to stereotypes of the elderly. Participants that were primed with an elderly stereotype walked down the hallway after leaving the experiment more slowly than those who were not primed with the stereotype. In another experiment, participants were primed with words relating to rudeness or politeness. Those primed with rudeness interrupted significantly more during a conversation task than those primed with politeness. Bargh et al. argued that an automatic activation of perceptual representations associated with stereotypes and trait constructs continuously activated behaviors that were consistent with these representations, which increased the likelihood that the individual produced the behaviors. In regards to our findings, listeners may have activated a set of general perceptual or perceptual motor representations related to either old or young age that in turn affected their execution of motor behavior in general, including their degree of accommodation in the shadowing task.

The findings from Experiment 2 provide evidence against this possibility. If participants' response latencies, in addition to the characteristics of their shadowing responses, reflected expectations about talker's age so that participants initiated responses more slowly when expecting utterances to be spoken by an older talker, and initiated responses more quickly when expecting utterances to be spoken by a younger talker, it would suggest that participants generally embodied stereotypes or expectations related to either young or old age and that these effects were not specific to spoken language and vocal accommodation. However, although participants' vocal accommodation was influenced by expected age of the speaker with shadowed utterances differing in duration as a function of age, participants' response latencies did not differ as a function of expected age of the talker. If participants embodied a general motor behavior related to age stereotype, then not only would their accommodation be influenced by expected age but also their time to initiate a response would be influenced by expected age. That response latencies were not similarly affected by age suggests that the primed expectations of age were specific to vocal accommodation.

Another possibility for how social variables may have influenced perception production links is that listeners' expectations about the vocal behavior of the talker specifically affected their *perception* of what they heard. That is, when listeners had an expectation that what they were about to hear was produced by an old speaker, listeners actually perceived these utterances as being slow rather than neutral in speech rate and thus accommodated towards this biased perception. Likewise, when listeners had an expectation that what they were about to hear was produced by a young speaker, listeners may have perceived these utterances to be faster in speech rate and accommodated towards an illusory perception of fast speech rate. In either case, on this account, social variables or expectations interacted with perceptual-motor processing by influencing perception. Others have found that social expectations or stereotypes about a talker (i.e. gender) can have an effect on listeners' perception of speech produced by the talker (Johnson, 2006; Johnson, Strand, & Imperio, 1999, Staum & Casasanto, 2008). Johnson et al. found that when listeners were told an ambiguous gendered speaker was either female or male, listeners' judgments of vowel identity differed based on gender expectations even though listeners were presented with the same acoustic stimuli, suggesting an influence of social expectations on perceptual processing of speech. To the extent that the time to initiate a naming response reflects perceptual processing, listeners were able to perceptually process the details from the acoustics inherent in the speech signal but produced a behavior that was inconsistent with what was being perceptually processed. This is not to say that perception is not tied to production. Rather, activating social expectations or stereotypes may have influenced the automatic connections between perception and production, leading to the production of a motor behavior consistent with the social expectation that was primed.

The shadowing task in Experiment 2 in which social context was minimal also reinforces that vocal accommodation does not rely solely on conversational or social interaction between interlocutors. Although the results demonstrated that social variables interacted with perception production links to influence the degree of accommodation, it does not undermine the notion that perception production links underlie accommodation, but rather lends support to the proposal that social variables and perception production links interact to contribute to the degree of vocal accommodation. That is, these perception production links could continuously be operating both within and outside a social domain during accommodation.

In context where social variables are minimal and social goals are unclear, any coupling between perception and production may operate relatively automatically as shown from the findings in Experiment 1b. However, even when interlocutors are interacting and there are social goals to be attained, the process linking perception and production may still occur automatically. In this case, an individual may need to make an explicit decision to continue to accommodate or to maintain their own vocal behavior. That is, the process of perceiving a vocal behavior, which then facilitates the production of a similar behavior, may be an automatic process resulting in a simulation or vocal representation that the individual automatically activates and then in some cases consciously chooses to inhibit. For example, during an interaction, depending on the desire to attain a social goal, the listener could choose to produce this simulated vocal behavior to become more similar in vocal style to their interlocutor or could inhibit this simulation to maintain his or her own vocal behavior or produce a vocal behavior that is dissimilar from the vocal characteristics of the talker.

In the case when listeners accommodate their vocal behavior based on a social expectation or stereotype rather than what is perceptually present in the acoustic signal, the social expectations or stereotypes themselves are often activated outside of conscious awareness (Bargh, 1994; 1999; Bargh & Ferguson, 2000; Chartrand, Maddux, & Lakin, 2005; Patterson, 2001). Some researchers have argued that individuals have social categories that are built of specific traits of members in that category (Hummert, 1990; Hummert, Shaner, & Garstka, 1995; Hummert, Shaner, Garstka, & Henry, 1998), and that stereotype formation comes from activating certain traits within these categories. This occurs from a set of impressions experienced over time so that when a behavior, trait, or cue is present in a situation, a judgment is made automatically about the individual with little effort. This may be one way listeners simulate a stereotypic vocal behavior. Simulations of behavior can reactivate perceptual states that were previously experienced as a result of interacting with the environment (Hesslow, 2002) and these simulations can involve perceptual information about the social context or situation. Representations of situation specific strategies can be stored in memory and invoked with little cognitive effort when a cue is presented that is similar to that previous experience. For instance, during the perception of a talker's vocal productions, characteristics of the talker's utterances are encoded and stored in memory. After certain vocal behaviors are associated with a social category and stored in an individual's motor repertoire, cues that are introduced in the situation can prime these prior perceptual experiences, which can then predispose individual to accommodate based on the stereotypic vocal behaviors rather than the observed vocal behavior. Thus an individual may have many experiences with interacting with various older talkers encoding characteristics about not only talker's voice related to this age group, but also the context or social interaction taking place. All of these instances are encoded into memory. When a cue in the environment is similar to what was previously experienced or expectations about the talker are introduced to the listener, these prior representations of vocal behavior can be activated, predisposing the listener to execute the stereotypic behavior related to that social category rather than what is perceptually present in the acoustic signal. Thus listeners simulate a stereotypic vocal behavior related to a social category rather than activating a vocal behavior consistent with the acoustic input itself.

This proposal is consistent with exemplar based approaches to speech perception (Goldinger, 1996, 1998; Johnson, 1997, 2005; Johnson, Strand, & Imperio, 1999; Pierrehumbert, 2001, 2002) suggesting that detailed exemplars of talker utterances which include talker characteristics are stored in memory and activated during speech perception. Johnson (2005) in particular has suggested that when perceiving spoken utterances associated with socio-variation, the type of representations involved could be explained using an exemplar-resonance model. Johnson proposed that the similarity between the acoustic input and detailed exemplars that are stored in memory by the listener determine the activation or resonance in the model. The strength of connection weights or resonance between exemplars and cognitive categories can account for the categorization of stimuli associated with socio-variation. Thus, the more similar the input is to the exemplar, the more category activation there is which leads to stimuli being placed into the category that has the greatest resonance. In the current experiment, because the input itself was ambiguous, the invoked stereotype influenced listeners' expectations about whether the speech stimuli were spoken by a young or old talker, which may have may have either led listeners to attend to or infer particular social category consistent features and then to place this ambiguous stimuli into a social category of exemplars with respect to age

. An alternative is that listener's experience with different talkers results in a normalized abstract representation (Joos, 1948; Ladefoged & Broadbent, 1957; Pisoni, 1997) that can characterize different categories of social variation. Indeed others have proposed theories of abstract representations for speech perception (McQueen, Cutler, & Norris, 2006; Lahiri & Marslen-Wilson, 1991) but have for the most part accounted for lexical variation rather than socio-variation among talkers. It may be that collections of exemplars related to a particular social category could result in summary representations that account for socio-variation in talker's voice rather than individual exemplars. Thus, an abstract or summary representation for characteristics of talker's voice rather than specific instances with particular characteristics of talker's voice, are represented in memory. This summary representation can represent properties of talker's voice that fall under some social category. For instance, a stereotypic vocal behavior could be a result of an abstract representation that the listener has formed from detailed exemplars of vocal utterances collected over time with talkers that are part of that stereotyped category. What a listener activates in terms of stereotypic vocal behavior is highly influenced by his or her experience with talkers that fall into that stereotypic social category. In either case, an exemplar based account, or an account that uses exemplars to form summary representations, experience with particular types of talker variation influence representation and processing.

With respect to the current findings, listeners in our study may have encoded detailed exemplars of utterances in memory both older talkers and younger talkers and either maintained that exemplar-based representation or formed an abstract representation of talker's voice related to a social category of age (either old or young). When listeners were primed with an expectation about a vocal behavior related to age, this could have activated collections of exemplars or distinct abstract vocal representation of age that then affected the degree of accommodation that occurred in response to the ambiguous acoustic stimuli. For example, when a listener was primed to expect a "young" talker, this expected vocal behavior activated exemplars that were associated with young talkers' utterances, which then resonated with a category of young age. Alternatively, listeners may have accessed an abstract vocal representation of young age that in turn affected their subsequent production. Likewise, when a listener was primed to expect an "old" talker, the expected vocal behavior activated exemplars or summary abstract representations that were associated with old talkers' utterances, which then resonated with a category of old age. These representations of old age would have included a slower speech rate, resulting in changes to their vocal accommodation.

Although an exemplar based resonance model is generally discussed as a perceptual account, this model could offer a similar account for how simulations of vocal behavior are activated and produced during accommodation. For example, when listeners were primed with an expectation about a vocal behavior related to age, this could have activated collections of previous simulations of motor behavior related to age that then affected the degree of accommodation that occurred in response to the ambiguous acoustic stimuli. For example, when a listener was primed to expect a "young" talker, this expected vocal behavior activated a motor simulation related to young talkers' voice, which then resonated with a category of young age. Alternatively, listeners may have accessed an abstract simulation of motor behavior based on young age that in turn affected their subsequent production. Likewise, when a listener was primed to expect an "old" talker, the expected vocal behavior activated motor simulations that were associated with old talkers' utterances, which then resonated with a category of old age.

Although an exemplar based resonance model could account for the type of representations being used or the type of motor simulations activated when social context interacts with perception production links during accommodation, it is important to note that on this account, it is listener's experience with various talkers that is responsible for forming these types of representations and simulations. This type of account would suggest that listeners would start by encoding specific representations of talker's voice into memory and then categorize these exemplars into different social domains to account for socio-variation among talkers' voices. This process would also presumably allow the listener to then form an abstract representation of talker's voice that is representative of some social category. Thus if we had listeners in our studies that had a very different experience with talking with old talkers so that their experience was that old talkers did not talk more slowly than young talkers, we may have observed that our manipulation of priming listeners with "old" may not have had such a significant effect on accommodation. That is, if listeners had many exemplars for which older talkers did not speak more slowly than young talkers, their resulting representations for older talkers' voices would of course not include a slower speech rate and listeners would not be expected to change their accommodation behavior or simulate a vocal behavior that included a slower speech rate. Changes in vocal behavior would only occur if they were consistent with previous experience.

A significant implication of these findings is that the social and cognitive mechanisms that have been proposed to underlie accommodation are not independent. These studies examined the contributions of both social and cognitive processes underlying accommodation and these results may begin to inform us about the processes and representations being used to support vocal accommodation. In addition, by examining these processes, we can begin to address how listeners are able to adjust their own vocal repertoire to become more similar to whom they are listening and what kind of representations might be implicated during accommodation. One way to address how listeners are able to vocally accommodate is to investigate the developmental precursors that may contribute to accommodation. For example, perception production links may operate automatically until learned social factors begin to influence either the nature of or the automaticity of this kind of perceptual-motor coupling. Perception production links may be activated at a very early age. Infants have been observed to imitate gestural behaviors hours after birth demonstrating a predisposition to imitate the actions of other individuals (Meltzoff & Moore, 1992, 1997) suggesting that there may be low level perceptual motor links that are already operating for some general motor behaviors. In relation to vocal accommodation and speech, infants only several days old are perceptually sensitive to their mother’s voice versus other female voices (DeCasper & Fifer, 1980) and become aware of the correspondences between certain kinds of facial and vocal activity by 3 or 4 months (Kuhl & Meltzoff, 1982; MacKain, Studdert-Kennedy, Speker, & Stern, 1983). As a result, a general link between perception and production in speech may be in place from birth, but the degree of accommodation that occurs may rely on the infant's experience with their own vocal articulatory movements at around 6-12 months of age (Imada, Zhang, Cheour, Taulue, Ahonene, & Kuhl, 2006). Simply perceiving or observing a motor action has been shown to facilitate the action of the behavior (Black & Wright, 2000; Heyes & Foster, 2002; Mattar & Gribble, 2005; Vinter & Perruchet, 2002; Vogt, 1995) suggesting that there may be little learning in terms of a perception production link underlying motor behavior. However, with speech production and communication, individuals must have experience with the production of the particular sounds of their language through their own vocal repertoire in order to simulate and produce sounds that are similar to the vocal characteristics of other talkers.

In addition, learning social constructs or gaining social experience influences these perception production links throughout development. However at younger ages, children may not have had enough social experience to recognize various social variables such as status differentiations and social norms to influence the automaticity of perceptual motor mechanisms that underlie accommodation of vocal behaviors. It may be that a general perception production link predisposes young children to accommodate their vocal behavior automatically and the degree to which social factors influence this automaticity may develop as children gain social experience.

The experiments presented here show how social variables can have a top down effect on perception production links during accommodation. In addition, these findings demonstrate how both social and cognitive accounts can interact as possible mechanisms underlying vocal accommodation. The notion that general perception production links can interact with social variables during accommodation has broader implications for spoken language processing in general. During speech, listeners face an enormous amount of variation among different talkers. One of the largest questions in speech perception research is how listeners contend with this variability in order to reach a meaningful linguistic percept. Some of this variation is due to the social aspects of the talker, which can not only influence perception of speech, but also can influence how the listener produces his or her own utterances during conversation. This in turn can influence the relationship and dynamics between interlocutors and the flow and intelligibility of the interaction. By understanding how social variables can influence perception production links underlying speech, we can begin to further address how listeners use this social variation when processing speech and the types of representations listeners are using when processing and responding to the social variation during communicative interaction with another talker.

The extent to which vocal accommodation involves both social and cognitive mechanisms suggests that spoken language between interlocutors is a highly dynamic system. That is, the degree of accommodation that occurs between interlocutors is not only motivated by the automaticity of relatively hard-wired perceptual motor links, but is also influenced by several social factors such as the built up social experiences that each interlocutor brings to the interaction, the social expectations interlocutors have of each other, the social relationship between interlocutors, and the desire to attain social goals on each part of the individual. Thus the exchange of spoken language between interlocutors is a complex system that is constantly changing as a result of an interaction between these social and cognitive mechanisms underlying behavior.

References

Abercrombie, D. (1967). *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

Adank, P., Hagoort, P., & Bekkering, H. (2010). Imitation improves language comprehension. *Association for Psychological Science*, 1903-1909.

Allen, J. S., and Miller, J. L. (2004). Listener sensitivity to individual talker differences in voice-onset-time. *Journal of the Acoustical Society of America*, *115,* 3171-3183.

Allen, J. S., & Miller, J. L. (1999). Effects of syllable-initial voicing and speaking rate on the temporal characteristics of monosyllabic words. *Journal of Acoustical Society of America, 106*, 2031-2039.

Amerman, J. D. & Parnell, M. M. (1992) Speech timing strategies in elderly adults, *Journal of Phonetics, 20,* 65-76.

Amerman J.D., Parnell M.M. (1990). Auditory impressions of the speech of normal elderly adults. *British Journal of Disorders of Communication*, *25,* 35–43.

Argyle, M., Furnham, A., & Graham, J.A. (1981). *Social Situations.* Cambridge: Cambridge University Press.

Babel, M. (2009). *Phonetic and Social Selectivity in Speech Accommodation.* Doctoral dissertation, Department of Linguistics, UC Berkeley.

Bargh, J. A. (1999). The cognitive monster: The case against controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology*. New York: Guilford.

Bargh, J. A. (1994). The Four Horsemen of automaticity: Awareness, efficiency, intention, and control in social cognition. In R. S. Wyer, Jr., & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., pp. 1-40). Hillsdale, NJ: Erlbaum.

Bargh, J.A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology, 71*, 230-244.

Bargh, J. A., & Ferguson, M. L. (2000). Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin, 126,* 925-945

Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P. and Pike, B. (2000) Voice-selective areas in human auditory cortex. *Nature*, 403, 309-312.

Berger, C.R. & Bradac, J.J. (1982). *Language and social knowledge*. London: Edward Arnold.

Bilous, F. & Krauss, R. M. (1988). Dominance and accommodation in the conversational behavior of same-and mixed-gender dyads. *Language and Communication, 8*, 183-194.

Binder, J.R., Frost, J.A., Hammeke, T.A., Bellgowan, P.S., Springer, J.A., & Kaufman, J.N. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex, 10,* 512-520.

Black, C.B., & Wright, D.L. (2000). Can observational practice facilitate error recognition and movement production? *Research Quarterly for Exercise and Sport, 71 (4),* 331-339.

Bookheimer, S.Y. (2002). Functional MRI of Language: new approaches to understanding the cortical organization of semantic processing. *Annual Review of Neurosciences, 25,* 151-158.

Bourhis, R. Y. & Giles, H. (1977). The language of intergroup distinctiveness. In H. Giles (Ed.), *Language, Ethnicity and Intergroup Relations* (pp. 119-135).London: Academic Press.

Bourhis, & Giles, H. (1976). The language of co-operation in Wales. *Language Sciences, 42*, 13-16.

Boursma, Paul, & Weenink, David. (2008). Praat: doing phonetics by computer (Version 5.0.09) [Computer program]. Retrieved September 4, 2009 from http://www.praat.org.

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B. & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Perception & Psychophysics, 61* (5), 977-985.

Bradlow, A.R., Pisoni, D. B., R. A. & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America, 101 (4),* 2299-2310.

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.

Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psycholigica,106,* 3–22.

Brass, M., Bekkering, H., Wohlschla¨ger, A., & Prinz, W. (2000). Compatibility between observed and executed finger movements: Comparing symbolic, spatial, and imitative cues. *Brain and Cognition, 44*, 124–143.

Brewer, M.B. (1988). A dual process model of impression formation. In T. K. Srull & R.S. (eds.). *Advances in Social Cognition*. Lawrence, Erlbaum Associates, Hillsdale, NJ, Vol. 1, (pp. 1-36).

Buccino G, Binkofski F, Fink GR, Fadiga L, Fogassi L, (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European Journal of Neuroscience, 13,* 400-404.

Byrne, D. (1971). *The Attraction Paradigm,* New York: Academic Press.

Caporael, L. R., and Culbertson, G. H. (1986). Verbal response modes of baby talk and other speech at institutions for the aged, *Language and Communication,6,* 99-11.

Cargile, A , Giles, H., & Clément, R. (1996). The role of language in ethnic conflict. *In* J. Gittler *(Ed.),* Racial and ethnic conflict: Perspectives from the social disciplines *(pp.* 189-208). Greenwich, CT: PAI Press.

Castiello, U., Lusher, M., Mari, M., Edwards, M., & Humphreys, G.W. (2002). Observing a human and a robotic hand grasping an object: Differential motor priming effects. In W. Prinz & B. Hommel (Eds.), *Attention and performance, Vol. XIX: Common mechanisms in perception and action* (pp. 315–333). New York: Oxford University Press.

Chartrand, T.L., & Bargh, J.A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology, 76,* 893–910.

Chartrand, T.L., Maddux, W., & Lakin, J. (2005). Beyond the perception-behavior link: The ubiquitous utility and motivational moderators of nonconscious mimicry. In R. Hassin, J. Uleman, & J.A. Bargh (Eds.), *Unintended thought II: The new unconscious*. (pp 334-361). New York: Oxford University Press.

Chen, M., & Bargh, J.A. (1997). Nonconscious behavioral confirmation processes: The Self-fulfilling consequences of automatic stereotype activation. *Journal of Experimental Social Psychology*, 33, 541-560.

Cho, T. & Ladefoged, P. (1999). Variations and universals in VOT: evidence from 18 languages. *Journal of Phonetics*, *27,* 207-229.

Chong TT, Cunnington R, Williams MA, Kanwisher N, Mattingley JB (2008). fMRI adaptation reveals mirror neurons in human inferior parietal cortex. *Current Biology, 18,* 1576 –1580.

Coats, E.J., & Feldman, R.S. (1996). Gender differences in nonverbal correlates of social status. *Personality and Social Psychology Bulletin, 22,* 1014-1022.

Coupland, N., Coupland, J., Giles, H., & Henwood, K. (1988). Accommodating the elderly: Invoking and extending a theory. *Language in Society, 17,* 1-41.

Coupland, N. & Giles, H. (Eds) (1988). Communication accommodation: Recent developments. *Language and Communication, 8,* 175-327.

DeCasper, A.J., & Fifer, W.P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science, 208*, 1174-1176.

Decety, J., Ingvar, D.H., (1990). Brain structures participating in mental simulation of motor behavior: a neuropsychological interpretation. *Acta Psychologica,. 73*, 13–24.

De Jong, R. (1993). Multiple bottlenecks in overlapping task performance. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 965-980.

De Jong, R., & Sweet, J.B. (1994). Preparatory strategies in overlapping-task performance. *Perception and Psychophysics*, *55*, 142-151.

Devine, P.G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56,* 680-690.

Dijksterhuis, A, & van Knippenberg, A. (1998). The relation between perception and behavior, or how to win a game of trivial pursuit. *Journal of Personality and Social Psychology*, 4, 865-877.

Dinstein, I., Hasson, U., Rubin, N., & Heeger, D.J. (2007). Brain areas selective for both observed and executed movements. *Journal of Neurophysiology*, 1415-1427.

Drost, U., Rieger, M., Brass, M., Gunter, T., & Prinz, W. (2003). *Instrument-specific action-effect associations in piano and guitar players.* Paper presented at the 44th annual meeting of the Psychonomic Society, Vancouver, British Columbia, Canada.

Duchin, S. W., & Mysak, E. D. (1987). Disfluency and rate characteristics of young adult, middle aged, and older males. *Journal of Communication Disorders, 20*, 245-257.

Fadiga L, Fogassi L, Pavesi G, & Rizzolatti G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of Neurophysiology*, *73,* 2608–2611.

Fecteau, S., Armony, J.L., Joanette, Y., & Belin, P. (2004). Is voice processing species-specific in human auditory cortex? An fMRI study, *Neuroimage, 23,* 840-848.

Feldstein, S. & Welkowitz, J. (1978). A chronography of conversation: In defense of an objective approach. In A.W. Siegman and S. Feldstein (Eds), *Nonverbal behavior and Communication*, (pp. 435-500). Hillsdale, NJ: Erlbaum.

Ferguson, M.J., & Bargh, J.A. (2004). How social perception can automatically influence behavior. *Trends in Cognitive Sciences*, 33-39.

Fernald, A., Taeschner, T., Dunn, J., & Papousek, M. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language, 16,* 477-501.

Ferrari, P.F., Rozzi, S., & Fogassi, L. (2005). Mirror neurons responding to observation of actions made with tools in monkey ventral premotor cortex. *Journal of Cognitive Neuroscience, 17*, 212-226.

Fowler, C. A. (2003). Speech production and perception. In A. Healy and R. Proctor (eds.). *Handbook of psychology, Vol. 4: Experimental Psychology*, (pp. 237-266). New York: John Wiley & Sons.

Fowler, C. A., Brown, J. M., Sabadini, L., & Weihing, J. (2003). Rapid access to speech gestures in perception: Evidence from choice and simple response time tasks. *Journal of Memory and Language, 49,* 396–413.

Fowler, C. A & Galantucci, B. (2005). The relation of speech perception and production. In D.B. Pisoni & R.E. Remez (Eds.). *The handbook of speech perception* (pp.633-652). Malden, MA: Blackwell.

Fowler, C. A., Galantucci, B. and Saltzman, E. (2003). Motor theories of perception. In M. Arbib (Ed.) *The handbook of brain theory and neural networks*. (pp. 705-707) Cambridge MA: MIT Press.

Fox, S.A., & Giles, H. (1996). Interability communication: Evaluating patronizing encounters. *Journal of Language and Social Psychology, 15*, 265-290.

Frick, R. W. (1985). Communicating emotion: The role of prosodic features. Psychological Bulletin, 97 (3), 412-429.

Galantucci, Fowler, C.A. & Turvey, M.T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin and Review*, *13* (3), 361–377.

Gallois, C. (2004). Communicating disability: Stereotypes, identity, and motivation. In S.H. Ng, C.H. Candlin, & C.Y. Chiu (Ed.), *Language matters: Communication, culture, and identity* (pp. 355-374). Honk Kong: City University of Hong Kong Press.

Gallois,C., & Callan,V.J. (1991). Interethnic accommodation: The role of norms. In H. Giles, J.Coupland, & N.Coupland (Eds.), *Contexts of accommodation: Developments in applied sociolinguistics* (pp.245-271).New York: Cambridge University Press.

Gallois, C., Giles, H., Jones, E., Cargile, A.C. & Ota, H. (1995). Accommodating intercultural encounters: Elaborations and extensions. In R. Wiseman (Ed.), *Intercultural communication theory* (pp. 115-147). Thousand Oaks, CA: Sage.

Gilbert, D. T., Krull, D. S. & Pelham, B. W. (1988). Of thoughts unspoken: Social inference and the self-regulation of behavior. *Journal of Personality and Social Psychology*, *55*, 685-694.

Giles, H. (1984). The dynamics of speech accommodation. *International Journal of the Sociology of Language, 46*, 1-155.

Giles, H. (1977). Social psychology and applied linguistics: Towards an integrative approach. *ITL Review of Applied Linguistics, 35*, 27-42.

Giles, H. (1973). Accent mobility: A model and some data. *Anthropological Linguistics*, *15*, 87-109.

Giles, H. & Coupland, N. (1991). Language attitudes: Discursive, contextual and gerontological considerations. In A.G. Reynolds (ed.), *Bilingualism, Multiculturalism, and Second Language Learning:The McGill Conference in Honor of Wallace E. Lambert*, (pp. 21-42). Hillsdale, NJ, Erlbaum.

Giles, H., Coupland, J., & Coupland, N. (1991). *Accommodation theory: Communication, context, and consequence*. In H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of accommodation: Developments in applied sociolinguistics* (pp. 1-68). New York: Cambridge University Press.

Giles, H., Coupland, N., Coupland, J., Williams, A., Nussbaum, J. (1992). Intergenerational talk and communication with older people. *International Journal of Aging and Human Development, 34*, 271-297.

Giles, H., Mulac, A., Bradac, J. J., & Johnson, P. (1987). Speech accommodation theory: The first decade and beyond. In M. McLaughlin (Ed.), Communication Yearbook. 10 (pp. 13-48). Beverly HUls, CA: Sage.

Giles, H. & Noels, K. (1997). Communication accommodation in intercultural encounters. In J. Martin, T. Nakayama, & L. Flores (Eds.*), Readings in cultural contexts* (pp. 139-149). Mountain View, CA: Mayfield.

Giles, H. & Powesland, P.F. (1975). *Speech style and social evaluation.* London: Academic Press.

Giles, H. & Robinson, W.P. (Eds.) (1990). *Handbook of Language and Social Psychology*. Chichester: John Wiley and Sons Ltd.

Giles, H., Scherer, K. R., & Taylor, D. M. (1979). Speech markers in social interaction. In H. Giles & K. R. Scherer (Eds.), *Social markers in speech* (pp. 343-381). Cambridge, Eng: Cambridge University Press.

Giles, H. & Smith, P.M. (1979). Accommodation theory: Optimal level of convergence. In H.Giles, & R. St. Clair (Eds.), *Language and Social Psychology*, (pp 45-65). Oxford: Blackwell.

Giles, H., Taylor, D.M. & Bourhis, R. (1973). Towards a theory of interpersonal accommodation through language: some Canadian data. *Language in Society* 2, 177–192.

Giles, H. , & Wadleigh, P.M. (1999). Accommodating nonverbally. In L.K. Guerrero, J.A. DeVito, & M.L. Hecht (Eds), *The nonverbal communication reader: Classic and contemporary readings* (2nd. ed., pp 103-161). Thousand Oaks, CA: Sage.

Goldinger, S.D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105*, 251-279.

Goldinger, S.D., & Azuma, T. (2004). Episodic memory reflected in printed word naming. *Psychonomic Bulletin & Review, 11*, 716-722.

Gorham-Rowan, M. M., & Laures-Gore, J. (2006). Acoustic-perceptual correlates of voice quality in elderly men and women. *Journal of Communication Disorders, 39*, 171–184.

Gregory, S. (1994). Sounds of power and deference: Acoustic analysis of macro social constraints on micro interaction. *Sociological Perspective*, *37*, 497–526.

Gregory, S. W. (1990). Analysis of fundamental frequency reveals covariation in interview partners' speech. *Journal of Nonverbal Behavior*, *14*, 237–251.

Gregory, S. W., Dagan, K. & Webster, S. (1997). Evaluating the relations of vocal accommodation in conversation partners' fundamental frequencies to perceptions of communication quality, *Journal of Nonverbal Behavior, 21*, 23-43.

Gregory, S. W., Green, B.E., Carrothers, R. M. & Dagan, K. A. (2001). Verifying the primacy of voice fundamental frequency in social status accommodation. *Language and Communication*, *21*, 37–60.

Gregory, S.W., & Hoyt, B.R. (1982.) Conversation partner mutual adaptation as demonstrated by fourier series analysis. *Journal of Psycholinguistic Research*, 35-46.

Gregory, S. W. Jr., & Webster, S. (1996). A nonverbal signal in voices of interview partners effectively predicts communication accommodation and social status perceptions. *Journal of Personality and Social Psychology*, *70*, 1231-1240.

Grosjean, M., Zwickel, J., & Prinz, W. (2009). Acting while perceiving: Assimilation precedes contrast. *Psychological Research, 73(1),* 3−13.

Hamilton, A., Wolpert, D., & Frith, U. (2004). Your own action influences how you perceive another person's action. *Current Biology, 14(6),* 493−498.

Hari R., Forss N., Avikainen S., Kirveskari S., Salenius S., & Rizzolatti G. (1998). Activation of human primary motor cortex during action observation: a neuromagnetic study. *Proceedings of the National Academy of Sciences,* 95, 15061–15065.

Hartman, D.E. & Danhauer, J.L. (1976): Perceptual features of speech for males in four perceived age decades. *Journal of the Acoustical Society of America*, 59(3), 713-715.

Hesslow, G. (2002). Conscious thoughts as simulation of behavior and perception. *Trends in Cognitive Sciences, 6*, 242-247.

Hewstone, M., & Giles, H. (1986). Social groups and social stereotypes in intergroup communication: A review and model of intergroup communication breakdown. In W. Gudykunst (Ed.), *Intergroup communication* (pp. 10-26). London: Edward Allen.

Heyes, C.M. & Foster, C.L. (2002). Motor learning by observation: Evidence from a serial reaction time task. *The Quarterly Journal of Experimental Psychology, 55A (2),* 593-607.

Hickok, G. (2001). Functional anatomy of speech perception and speech production: psycholinguistic implications. *Journal of Psycholinguistic Research*, *30*, 225 – 234.

Hickok, G., Buchsbaum, B. Humphries, C. & Muftuler, T. (2003). Auditory-motor interaction revealed by fMRI: Speech, music, and working memory in area spt. *Journal of Cognitive Neuroscience*, *15,* 673-682.

Hummert, M.L. (1999). Age stereotyping: Social cognitive and developmental issues. In T. M. Hess & F. Blanchard-Fields (Eds.), *Social cognition and aging* (pp. 175-196). New York: Academic Press.

Hummert, M.L. (1994). Physiognomic cues and the activation of stereotypes of the elderly in interaction. *International Journal of Aging and Human Development, 39*, 5-20.

Hummert, M.L. (1990). Multiple stereotypes of elderly and young adults: A comparison of structure and evaluations. *Psychology and Aging, 5,* 182 193.

Hummert, M.L., Garstka, T.A., Ryan, E.B., & Bonnesen, J.L. (2004).The role of age stereotypes in interpersonal communication. In J. F. Nussbaum and J. Coupland (Eds.). *Handbook of communication and aging research* (2nd ed., pp. 91-121). Mahwah, NJ: Lawrence Erlbaum.

Hummert, M.L., Garstka, T.A., & Shaner, J.L. (1997). Stereotyping of older adults: The role of target facial cues and perceiver characteristics. *Psychology and Aging, 12*, 107-114.

Hummert, M.L., Mazloff, D., & Henry, C. (1999). Vocal characteristics of older adults and stereotyping. *Journal of Nonverbal Behavior*, *23*, 111-132.

Hummert, M.L., Shaner, J.L. Garstka, T.A. & Henry, C. (1998). Communication with older adults: The influence of age stereotypes, context, and communicator age. *Human Communication Research*, *25*, 125-152.

Iacoboni M., Woods R.P., Brass M., Bekkering H., Mazziotta J.C., Rizzolatti G. (1999). Cortical mechanisms of human imitation. *Science 286*, 2526-2528.

Imada T., Zhang Y., Cheour M., Taulue S., Ahonene A. & Kuhl, P. K. (2006). Infant speech perception activates Broca's area: A developmental magnetoencephalography study. *NeuroReport****.*** *17*, 957-962.

Ito, T., & Ostry, D.J. (2010). Somatosensory contribution to motor learning due to facial skin deformation. *Journal of Neurophysiology*, 1230-2010.

Jacoby, L. L. (1983). Perceptual enhancement: Persistent effects of an experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 21-38.

Johnson, K. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics, 34*, 485-499.

Johnson, K. (2005). Speaker normalization in speech perception. In Pisoni, D.B. & Remez, R. (eds) *The Handbook of Speech Perception*. Oxford: Blackwell Publishers (pp. 363-389).

Johnson, K. (1997). Speech perception without speaker normalization: An exemplar model. In Johnson & Mullennix (eds) *Talker Variability in Speech Processing.* San Diego: Academic Press (pp. 145-165).

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

Joos, M. (1948). Acoustic Phonetics. *Language Monograph, 23.* Baltimore, MD: Linguistic Society of America.

Kemp, W., & Yaeger-Dror, M. (1991). Changing realizations of A in -*(a)tion* in relation to the front A - back A opposition in Quebec French. In P. Eckert (ed.), *New Ways of Analyzing Sound Change*. New York: Academic Press, pp.127-184. (3.2.7)

Kemper, S. (1994). Elderspeak: Speech Accommodations to Older Adults: *Aging and Cognition, 1*, 17-28.

Kilner, J.M., Neal, A., Weiskopf, N., Friston, K.J., Frith, C.D. (2009). Evidence of mirror neurons in human inferior frontal gyrus. *The Journal of Neuroscience*, 10153-10159.

Kilner, J. M., Paulignan, Y., & Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Current Biology, 13*, 522–525.

Kohler, E., Keysers, C., Umilt`a, M.A., Fogassi, L., Gallese, V. & Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science, 297*, 846–848

Krauss, R. M. & Pardo, J. S. (2006). Speaker perception and social behavior: Bridging social psychology and speech science. In P.A.M. van Lange (Ed.), *Bridging Social Pyschology: The Benefits of Transdisciplinary Approaches* (pp 273-278). Hillsdale, NJ: Erlbaum.

Kucera, & Francis, W.N., (1967). *Computational Analysis of Present-day American Engish*. Providence: Brown University press.

Kuhl, P.K. & Meltzoff, A.N. (1982). The bimodal perception of speech in infancy. *Science, 218*, 1138-1141.

Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Ladefoged, P., & Broadbent, D. E. (1957). Information conveyed by vowels. *Journal of the Acoustical Society of America*, 29, 98-104.

Lahiri, A., & Marslen-Wilson, W.D. (1991). The mental representation of lexical form: A phonological approach to the mental lexicon. *Cognition*, *38*, 245-294.

Lawson-Sako, S. & Sachdev, I. (1996). Ethnolinguistic communication in Tunisian streets: Convergence and divergence. In Y. Suleiman (Ed.), *Language and identity in the Middle East and North Africa.* (pp. 61-79). Surrey: Curzon Press.

Liberman, P. (1967). *Intonation, perception, and language*. Cambridge, MA: The MIT Press.

Liberman, A., Cooper, F., Shankweiler, and M. Studdert-Kennedy (1967). Perception of the speech code. *Psychological Review*, *74*, 431-461.

Liberman, A. M., & Mattingly, I.G. *(*1985*)*. The motor theory of speech perception revised. *Cognition, 21*, 1-36

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., and Yamada, T. (1994). Training Japenese listeners to identify English /r/ and /l/: III. Long-term retention of new phonetic categories. *Journal of the Acoustical Society of America, 96,* 2076-2087.

Logan, J.S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America*, *89,* 874-886.

MacKain, K.S., Studdert-Kennedy, M., Spieker, S., & Stern, D. (1983). Infant intermodal speech perception is a left hemisphere function. *Science, 219,* 1347-1349.

Magnuson, J.S., & Nusbaum, H.C. (2007). Acoustic differences, listener expectations, and the perceptual accommodation of talker variability. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 391-409.

Martin, G.B., & Clark, R.D. (1982). Distress crying in neonates. Species and peer specificity. *Developmental Psychology*, *18*, 3-9.

Mattar, A.G., & Gribble, P.L. (2005). Motor learning by observing. *Neuron, 46(1),* 153-160.

Matarazzo, J. D. (1973). A speech interaction system. In D. J. Kiesler (Ed.), *The process* *of psychotherapy* (pp. 22–35). Chicago: Aldine.

McCann, R., Giles, H. (2006). Ageism and the workplace: A communication perspective. In Nelson, T. (Ed.), *Ageism* (pp. 163-199). Cambridge, Massachusetts: MIT Press.

McClelland, J.L., & Chappell, M. (1998). Familiarity breeds differentiation: A subjective-likelihood approach to the effects of experience in recognition memory. *Psychological Review, 105*, 724-760.

McQueen, J.M., Cutler, A., & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cognitive Science, 30(6)*, 1113-1126.

Meltzoff, A. N., & Moore, M. K. (1997). Explaining facial imitation: A theoretical model. *Early Development and Parenting, 6*, 179-192.

Meltzoff, A. N., & Moore, M. K. (1992). Early imitation within a functional framework: The importance of person identity, movement, and development. *Infant Behavior and Development, 15*, 479-505.

Miller, K. (2005). *Communication theories: Perspective, processes and contexts* (2nd ed). NY: McGraw Hill.

Miller, J.L., Grosjean, F., & Lomanto, C. (1984). Articulation rate and its variability in spontaneous speech: A reanalysis and some implications. *Phonetica, 41*, 215-225.

Minear M, Park DC. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers, 36*, 630–633.

Montepare, J. M., Steinberg, J., and Rosenberg, B. (1992). Characteristics of vocal communication between young adults and their parents and grandparents. *Communication Research, 19,* 479-492.

Mulac, A., & Giles, H. (1996). You're only as old as you sound: Perceived vocal age and social meanings. *Health Communication, 8,* 199-216.

Mullennix, J.W., & Pisoni, D.B. (1990). Stimulus variability and processing dependencies in speech perception. *Perception and Psychophysics, 47,* 379-390.

Mullennix, J.M., Pisoni, D.B., & Martin, C.S. (1989). Some effects of talker variability on spoken word recognition. *Journal of Acoustical Society of America*, *85,* 365-378.

Müsseler, J., & Hommel, B. (1997). Blindness to response-compatible stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 861-872

Müsseler, J. & Prinz, W. (1996). Action planning during the presentation of stimulus sequences: Effects of compatible and incompatible stimuli. *Psychological Research, 59,* 48-63.

Müsseler, J., Wuhr, P., & Prinz, W. (2000). Varying the response code in the blindness to response-compatible stimuli. *Visual Cognition, 7,* 743-767.

Namy, L. L., Nygaard, L.C., & Sauerteig, D. (2002). Gender differences in vocal accommodation: The role of perception. *Journal of Language and Social psychology, 21*, 422-432.

Natale, M. (1975a). Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality and Social Psychology, 32,* 790–804.

Natale, M. (1975b). Social desirability as related to convergence of temporal speech patterns. *Perceptual and Motor Skills, 40,* 827-839.

Niedenthal, P. M., Brauer, M., Halberstadt, J. B., & Innes-Ker, Å. H. (2001). When did her smile drop? Facial mimicry and the influences of emotional state on the detection of change in emotional expression. *Cognition & Emotion, 15,* 853– 864.

Nishitani N, Hari R. (2002). Viewing lip forms: Cortical dynamics. *Neuron 36,* 1211–1220.

Nishitani, N., & Hari, R. (2000). Temporal dynamics of cortical representation for action. *PNAS, 97*, 913-918.

Nygaard, L. C., Burt, S.A., & Queen, J.S. (2000). Surface form typicality and asymmetric transfer in episodic memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1228-1244.

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

Nygaard, L.C., Sommers, M., & Pisoni, D. (1994). Talker-specific learning in speech perception. *Perception and Psychophysics 60*, 355–376.

Özdemir, E., Norton, A., & Schlaug, G. (2006). Shared distinct neural correlates of singing and speaking. *Neuroimage, 33*, 628-635.

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, & Cognition*, *19*, 309-328.

Pardo, J.S. (2006). On phonetic convergence during conversational interaction. *Journal of Acoustical Society of America, 119* (4), 2382-2393.

Pardo, J.S & Remez, R.E. (2006). The perception of speech. In M. Traxler & M. Gernsbacher (Eds), *The Handbook of Psycholinguistics, Second Edition.* Cambridge, MA: Elsiver.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116*, 220-244.

Patterson, M.L. (2001). Toward a comprehensive model of nonverbal communication. In W. P. Robinson & H. Giles (Eds.), *The new handbook of language and social psychology* (pp. 159–176). New York: Wiley.

Pazzaglia, M., Pizzamiglio, L., Pes, E., & Aglioti, S.M. (2008). The sounds of action in apraxia. *Current Biology, 18,* 1766-1772.

Pierrehumbert, J. (2001) Exemplar dynamics: Word frequency, lenition, and contrast In J. Bybee and P. Hopper (eds.) *Frequency effects and the emergence of lexical structure*. John Benjamins, Amsterdam. 137-157.

Pickering, M. & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences,* 27(2), 169-226.

Pisoni, D. B. (1997). Some thoughts on "normalization" in speech perception. In K. Johnson & J.W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 9-32). San Diego: Academic Press.

Pisoni, D. B. (1993). Long-term memory in speech perception: Some new findings on talker variability, speaking rate and perceptual learning. *Speech Communication, 13*, 109-125.

Pisoni, D. B., & Martin, C. S. (1989). Effects of alcohol on the acoustic phonetic properties of speech: Perceptual and acoustic analysis. *Alcoholism: Clinical and*

Experimental Research, 13, 577-587.

Pratto, F., & Bargh, J. A. (1991). Stereotyping based on apparently individuating information: Trait and global components of sex stereotypes under attention overload. *Journal of Experimental Social Psychology*, *27,* 26-47.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology, 9*, 129-154.

Ptacek, P. H., & Sander, E. K. (1966). Age recognition from voice. *Journal of Speech and Hearing Research, 9,* 273-277.

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience, 6,* 576-582.

Putnam, W., & Street, R. (1984). The conception and perception of noncontent speech performance: Implications for speech accommodation theory. *International Journal of the Sociology of Language*. *46,* 97-114.

Ramig, L. A. (1983). Effect of physiological aging on vowel spectral noise. *Journal of Gerontology, 38*, 223–225.

Ramig, L. A., & Ringel, R. L. (1983). Effects of physiological aging on selected acoustic characteristics of voice. *Journal of Speech and Hearing Research, 26*, 22–30.

Remez R.E, Fellowes J.M., Rubin P.E. (1997) Talker identification based on phonetic information. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 651–666.

Rizzolatti, G. & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences, 21*, 188-194.

Rizzolatti G, & Craighero L. (2004). The mirror-neuron system. *Annual Review of*

Neuroscience, 27, 169-192.

Rizzolatti G, Fadiga L, Matelli M, Bettinardi V, & Paulesu E. (1996). Localization of grasp representation in humans by PET: 1. Observation versus execution. *Experiments in Brain Research, 111:* 246–52.

Rizzolatti, G., L. Fogassi. L. & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action, *Nature Review of Neuroscience*, 661–670.

Ryan, W. J. (1972). Acoustic aspects of the aging voice. *Journal of Gerontology, 27,* 265-268.

Ryan, W. J. & Burk, K. W. (1974). Perceptual and acoustic correlates of aging in the speech of males. *Journal of Communication Disorders,* 181-192.

Ryan, E. B., & Capadano, H. L. (1978). Age perceptions and evaluative reactions toward adult speakers. *Journal of Gerontology, 33,* 98-102.

Ryan, E., Giles, H., Bartolucci, G., & Henwood, K. (1986). Psycholinguistic and social psychological components of communication by and with the elderly. *Language and communication, 6,* 1-24.

Rubin, D. L. (1992). Non-language factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research In Higher Education, 33* (4), 511-531.

Scherer, K. R. (1979). Personality markers in speech. In K. R. Scherer and H. Giles (Eds.), *Social Markers in Speech.* Cambridge: Cambridge University Press.

Scherer, K. & Giles, H. (eds.). *Social Markers in Speech.* Cambridge: Cambridge University Press, 1979.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user's guide.* Pittsburgh: Psychology Software Tools Inc. (pp. A-19-A-35).

Schötz, S. (2007) Acoustic analysis of adult speaker age. In C. Müller (Ed.) *Speaker Classification I, Lecture Notes in Computer Science*, Springer, (pp 88-107).

Schötz, S. (2003). Towards synthesis of speaker age: A perceptual study with natural, synthesized, and resynthesized stimuli. *Phonum, 9,* 153-156.

Shepard, C. A., Giles, H., & Le Poire, B. A. (2001). Communication accommodation theory. In W. P. Robinson & H. Giles (Eds.), *The new handbook of language and social psychology* (pp. 33–56). New York: Wiley.

Shipp, T., Qi, Y., Huntley, R., & Hollien, H. (1992). Acoustic and temporal correlates of perceived age. *Journal of Voice*, *6*, 211–216.

Shockley, K., Sabadini, L., Fowler, C.A. (2004). Imitation in shadowing words. *Perception & Psychophysics, 66,* 422-429.

Skipper, J.I., Nusbaum, H.C., & Small, S.L. (2005). Listening to talking faces: Motor cortical activation during speech perception. *Neuroimage*, 25 (1), 76-89.

Smith, E. R. (1990). Content and process specificity in the effects of prior experiences. In R. S. Wyer & T. K. Srull (Eds.), *Advances in social cognition* (Vol. 3, pp. 1-60). Hillsdale, NJ: Erlbaum.

Staum Casasanto, L. (2008). Does Social Information Influence Sentence Processing? In B. C. Love, K. McRae, and V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Stewart, M. A., & Ryan, E. B. (1982). Attitudes toward younger and older adult speakers:

Effects of varying speech rates. *Journal of Language and Social Psychology,* 1, 91-109.

Street, R. L., Jr. (1984). Speech convergence and speech evaluation in fact-finding interviews. *Human Communication Research, 11*, 139-169.

Street, R. L., Jr. (1982). Evaluation of noncontent speech accommodation. *Language and Communication, 2*, 13-31..

Street, R. & Brady, R. (1982). Speech rate acceptance ranges as a function of evaluative domain, listener speech rate, and communication context. *Communication Monographs, 49*, 290-308.

Street, R., Jr., Brady, R., & Putnam, W. (1983). The influence of speech rate stereotypes and rate similarity on listener's evaluations of speakers. *Journal of Language and Social Psychology, 2,37*-56.

Street, R. L. and Giles, H. (1982). Speech accommodation theory: A social cognitive approach to language and speech behavior. In M. Roloff and C. R. Berger (eds.), *Social cognition and communication* (pp. 193-226). Beverly Hills, CA: Sage.

Sweeting, P. M., & Baken, R. J. (1982). Voice onset time in a normal-aged population. *Journal of Speech and Hearing Research, 25*, 129–134.

Thakerar, J. N. & Giles, H. (1981). They are-so they speak: Noncontent speech stereotypes. *Language and Communication, 1*, 251-260.

Thakerar, J. N. Giles, H. & Cheshire, J. (1982). Psychological and linguistic parameters of speech accommodation theory. In C. Fraser and K.R. Scherer (Eds.), *Advances in the Social Psychology of Language* (pp.205-255). Cambridge: Cambridge University Press.

Thelen, E., & Smith, L. B. (1994). A dynamic systems approach to the development of cognition and action. Cambridge, MA: MIT Press.

Triandis, H. (1960). Cognitive similarity and communication in a dyad. *Human Relations, 13,* 279-287.

Trudgill, P.J. (1986). *Dialects in Contact*. Oxford: Blackwell.

Tucker, M., & Ellis, R. (1998). On the relations of seen objects and components of potential actions. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 830–846.

Tucker, M., & Ellis, R. (2001). The potentiation of grasp types during visual object categorization. *Visual Cognition, 8*, 769–800.

Van Lancker, D., Kreiman, J., & Emmorey, K. (1985). Familiar voice recognition : Patterns and parameters. Part I. *Journal of Phonetics, 13,* 19-38.

Vinter, A. & Perruchet, P. (2002). Implicit motor Learning through observational training in adults and children. *Memory and Cognition, 30(2),* 256-261.

Vogt, S. (1995). On relations between perceiving, imagining and performing in the learning of cyclical movement sequences. *British Journal of Psychology, 86*, 191-216.

Webb, J. T. (1970). Interview synchrony: An investigation of two speech rate measures in an automated standardized interview. In A. W. Siegman & B. Pope (Eds.), *Studies in dyadic communication* (pp. 55–70). Oxford, NY: Pergamon.

Watkins, K. & Paus, T. (2004). Modulation of motor excitability during speech perception: The role of Broca's area. *Journal of Cognitive Neuroscience, 16*, 978-987.

Watkins, K.E., Strafella, A.P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*, *48*, 989-994.

Williams, A. & Giles, H. (1996). Intergenerational conversations: Young adults' retrospective accounts. *Human Communication Research*, *23*, 220-250.

Wilson, S.M, Ayse, P.S., Sereno, M.I., Iacoboni, M. (2004). Listening to speech activates motor areas in speech production. *Nature Neuroscience*, *7*, 701-702.

Wilson, M. & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, 131, 460-473.

Witt, Kemmerer, Linkenauger, & Culham (2010). A functional role for motor simulation in identifying tools. *Psychological Science*, 1-5.

Yonan, C. A. & Sommers, M. S. (2000). The effects of talker familiarity on spoken word identification in younger and older listeners. *Psychology and Aging*, *15*, 88-99.

Zatorre, R.J., Meyer, E., Gjedde, A., & Evans, A.C. (1996). PET studies of phonetic processing of speech: review, replication, and reanalysis. *Cerebral Cortex*, *6*, 21-20.

Zwaan, R. A., StanWeld, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science, 13*(2), 168–171.

Zwickel, J., Grosjean, M., & Prinz, W. (2010). On interference effects in concurrent perception and action. *Psychological Research*, *74*, 152−171.

Appendix A

Experimental Stimuli used for Experiment 1a, 1b, and 2

active	handle
badly	jelly
beacon	listen
beauty	mingle
buses	novel
canoe	nugget
careful	parcel
chicken	perfect
china	response
closely	rider
cousin	shimmer
cushion	sickness
dirty	staple
dozen	sugar
favor	title
fifty	twisted
follow	venom
forget	vision
frosty	wedlock
garden	winding

Appendix B

Descriptions of Old and Young Talkers in Experiment 2

Old Talker Description

This is Mr. Jones. He has been a participant in the speech perception lab in the past. He is a 70 year old male that has now retired to Florida. His skin is soft and wrinkly and his hair is mostly white with some grey undertones. Mr. Jones is not very modern in terms of fashion or lifestyle. He likes to wears argyle sweaters or cardigans and shuffles around in wool socks and slippers. He doesn't go out very often because he had replacement hip surgery last fall and so he is very cautious and careful whenever he walks somewhere. Mr. Jones is rather traditional and does not have internet at home. He doesn't believe in cell phones or computers. In fact, he finds newer technology and gadgets as more of a hassle than entertainment. He does not watch much tv. He prefers to write letters by hand. It takes him some time to finish them and put them in the mail. In turn, Mr. Jones appreciates receiving hand written mail. Mr. Jones enjoys playing Bingo once a week. He also enjoys it when his grandchildren come to visit him. Mr. Jones worries sometimes about living alone because he is often very forgetful. He is dependent on a watch alarm to remind him to take his arthritis medication in the mornings. A lot of times, Mr. Jones finds it difficult to go outside of the house because of his newly replaced hip. However, Mr. Jones is usually in good spirits and he enjoys trying to get out into the community and contribute as much as he can. For example, he has helped our projects in the lab by coming in to record speech stimuli before he moved to a retirement community in Orlando.

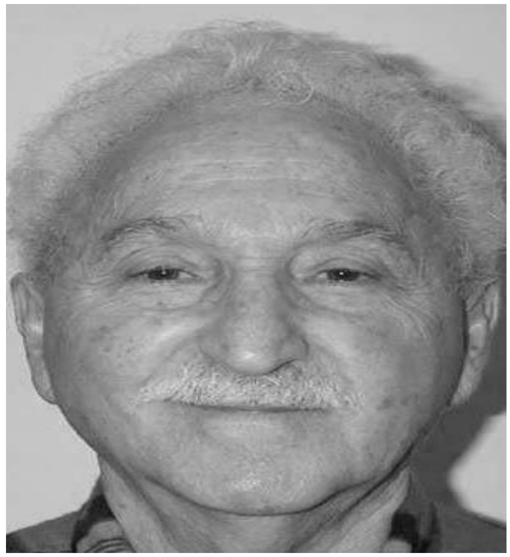
Young Talker Description

This is Tommy. Tommy has participated in our paid research studies. He is a 22 year old male that has moved from NY city. Although he was raised in NY, he has quickly adapted to Atlanta city life. Tommy is on a community rugby team for males 20-25 years of age and he plays at least once a week. Although Tommy is very athletic he does enjoy himself and likes to go out and party with his friends downtown. He prefers beer over liquor but will drink both. Tommy is very outgoing and is the first to get his group of friends pumped about doing something. For example, last spring break, Tommy coordinated a trip for him and four friends to go on a cruise to the Caribbean. Tommy is always on the go and doesn't sit around very much. Even at 22 he attends school, has a job as a waiter in a restaurant in Buckhead, and has to take care of his 5 year old lab. He is very independent and energetic and has lived away from home for over a few years. He is quite extroverted and is involved in both academic and extracurricular activities in the University. Tommy likes being the first one out of his group of friends to get the latest electronic gadgets. He bought the Iphone 4 when it came out and has the new mini iphone nano with a touch screen. Like most college students he has a twitter and facebook account. He is pretty addicted to being social and when he isn't hanging out with his friends he is either running with his dog at Piedmont park or riding his bike around. He isn't against having a girlfriend but is enjoying the single life.

Appendix C

Pictures of Old and Young Talkers in Experiment 2

Old Talker



Young Talker

Table 1
Average Durations and Standard Deviations of Stimulus Words Produced by each Old and Young Talkers at Fast and Slow Speaking Rates

	Age (years)	Fast Speaking Rate (ms)		Slow Speaking Rate (ms)	
Talkers		M	SD	M	SD
Old Male 1	67	398.05	41.63	890.40	46.53
Old Male 2	68	394.22	31.40	874.87	91.46
Young Male 1	23	394.02	42.89	902.22	66.88
Young Male 2	26	387.90	31.40	888.51	46.53

Table 2

Means and Standard Deviations for Proportion Correct Perceptual Identification of AGE as a Function of Age of the Speaker and Speaking Rate

	Age of Talkers					
	Old Talkers			Young Talkers		
Speaking Rate	OM1	OM2	M (Old)	YM1	YM2	M (Young)
Fast	0.68 (0.22)	0.31 (0.20)	0.49 (0.19)	0.86 (0.15)	0.89 (0.15)	0.88 (0.13)
Slow	0.94 (0.07)	0.77 (0.20)	0.85 (0.12)	0.77 (0.29)	0.67 (0.29)	0.72 (0.26)
M (Speaker)	0.81 (0.14)	0.54 (0.20)	0.67 (0.15)	0.81 (0.22)	0.78 (0.22)	0.80 (0.19)

Table 3
Means and Standard Deviations for Proportion Correct Perceptual Identification of SPEAKING RATE as a Function of Age of the Speaker and Speaking Rate

	Age of Talkers					
	Old Talkers			Young Talkers		
Speaking Rate	OM1	OM2	M (Old)	YM1	YM2	M (Young)
Fast	0.83 (0.13)	0.89 (0.13)	0.85 (0.13)	0.86 (0.13)	0.92 (0.11)	0.88 (0.12)
Slow	0.91 (0.12)	0.92 (0.12)	0.91 (0.12)	0.92 (0.12)	0.90 (0.12)	0.91(0.12)
M (Speaker)	0.87 (0.12)	0.90 (0.12)	0.88 (0.12)	0.46 (0.12)	0.91 (0.11)	0.98(0.12)

Figure Caption

Figure 1. *Identification Performance for Age as a Function of Age of the Speaker and Speaking Rate (Error bars represent standard error).*

Figure 2. *Identification Performance for Speech Rate as a Function of Age of the Speaker and Speaking Rate (Error bars represent standard error).*

Figure 3. *Age Identification Performance for Individual Old and Young Speakers as a Function of Speaking Rate (Error bars represent standard error).*

Figure 4. *Speech Rate Identification Performance for Individual Old and Young Speakers as a Function of Speaking Rate (Error bars represent standard error).*

Figure 5. *Average Difference Scores across Age and Speech Rate (Error bars represent standard error).*

Figure 6. *Average Response Latencies across Age and Speech Rate (Error bars represent standard error)*

Figure 7. *Average Difference Scores for Accommodation Performance Following Old versus Young Primes (Error bars represent standard error).*

Figure 8. *Average Response Latencies for Accommodation Performance Following Old versus Young Primes (Error bars represent standard error)*

Figure 1

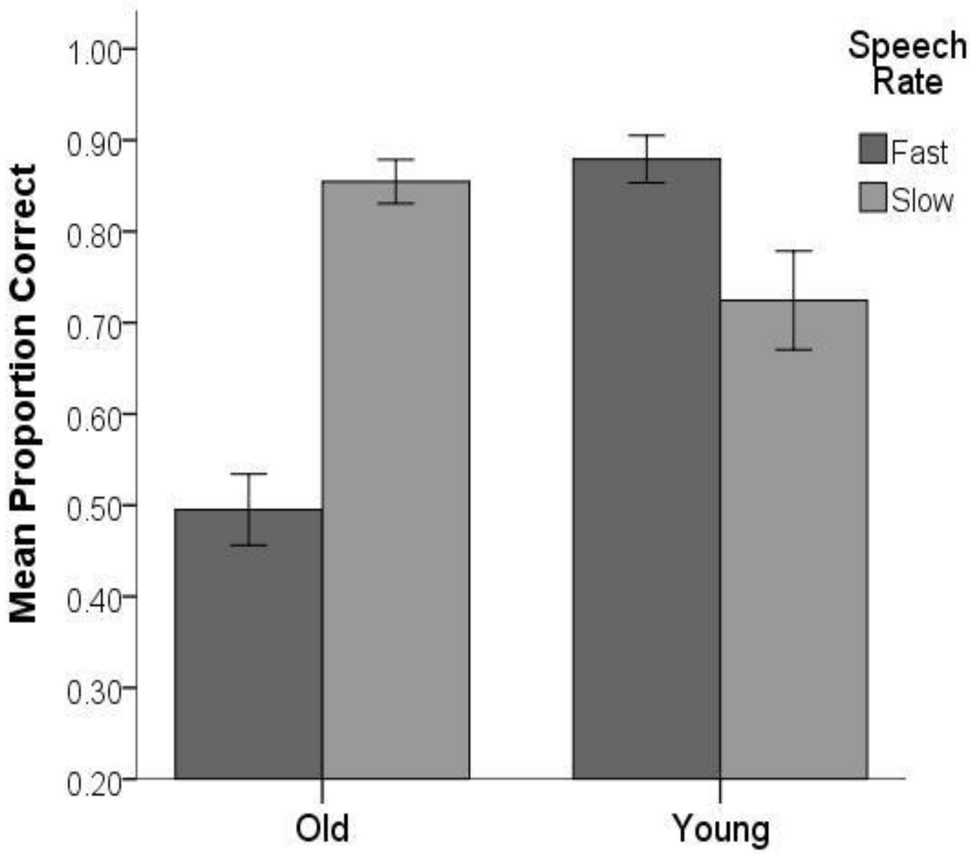

Mean Proportion Correct
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
Speech Rate
Fast
Slow
Old
Young

Figure 2

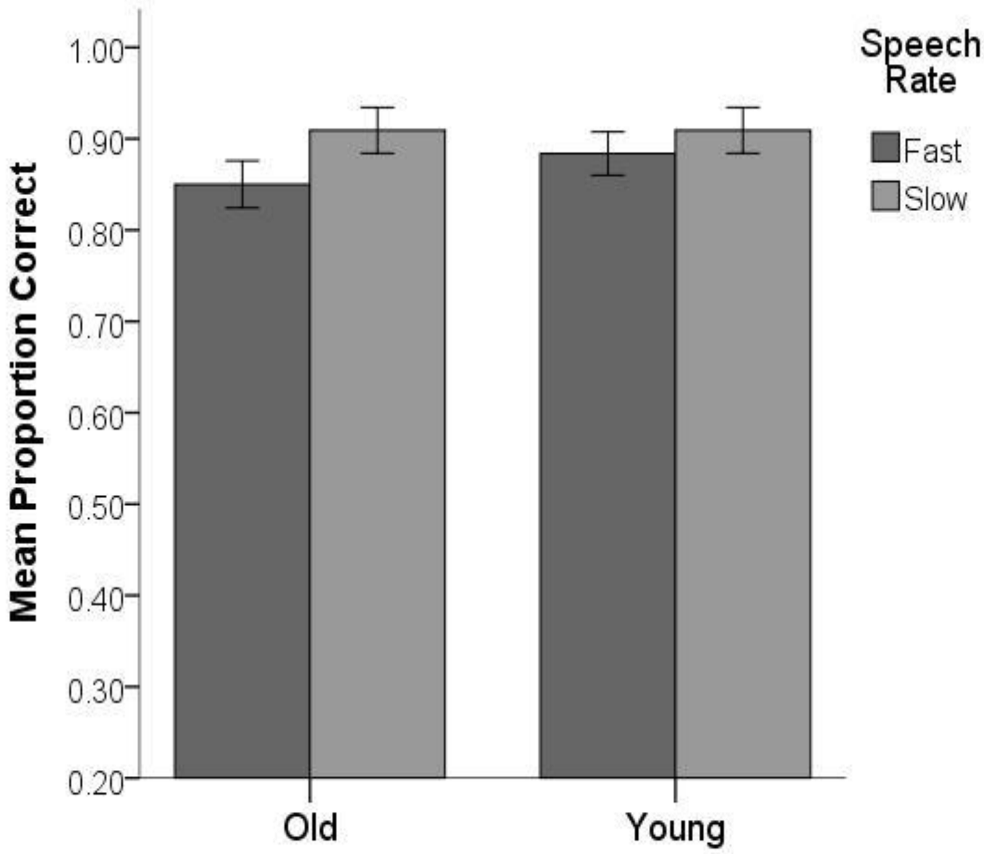

Mean Proportion Correct
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
Old
Young
Speech Rate
Fast
Slow

Figure 3

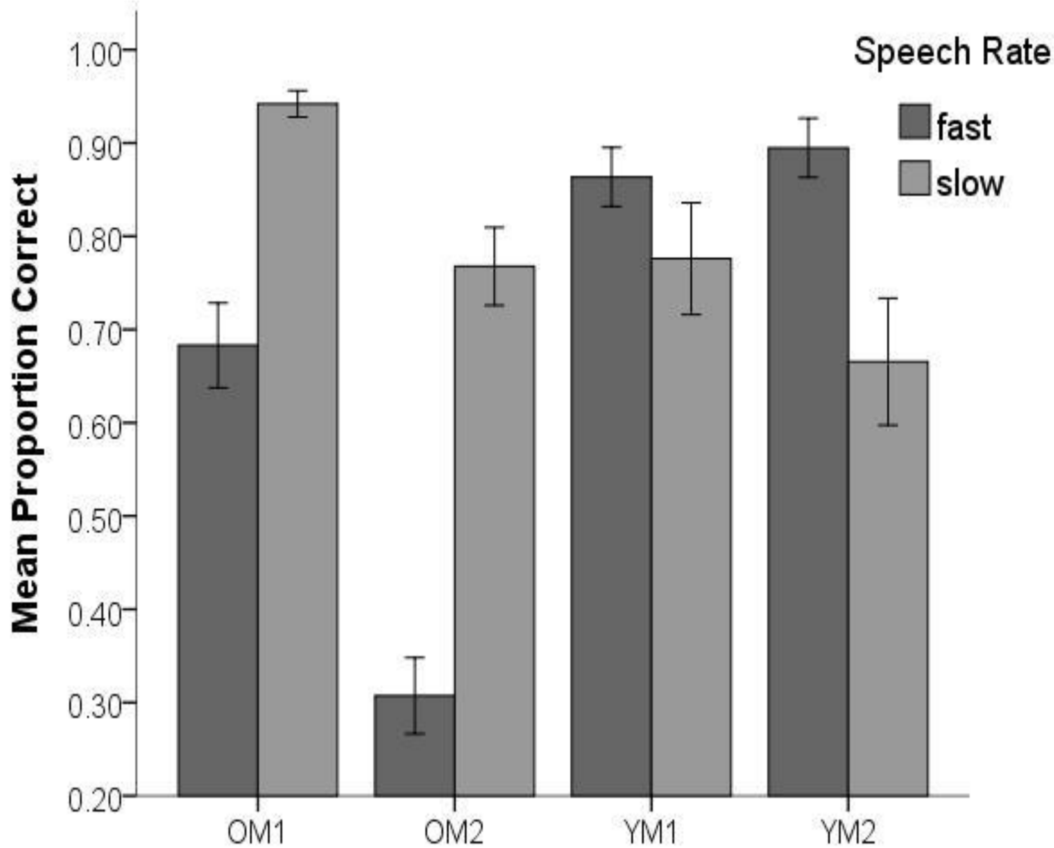

Speech Rate
fast
slow
Mean Proportion Correct
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
OM1
OM2
YM1
YM2

Figure 4

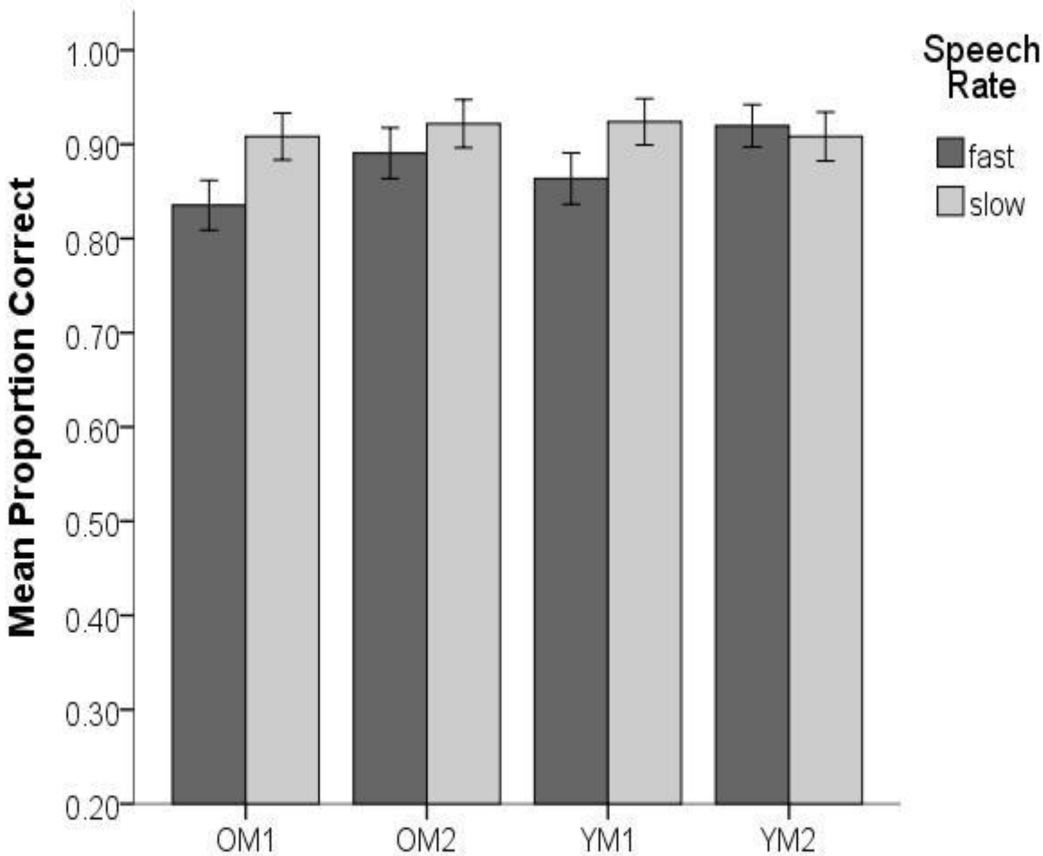

Mean Proportion Correct
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
OM1
OM2
YM1
YM2
Speech Rate
fast
slow

Figure 5

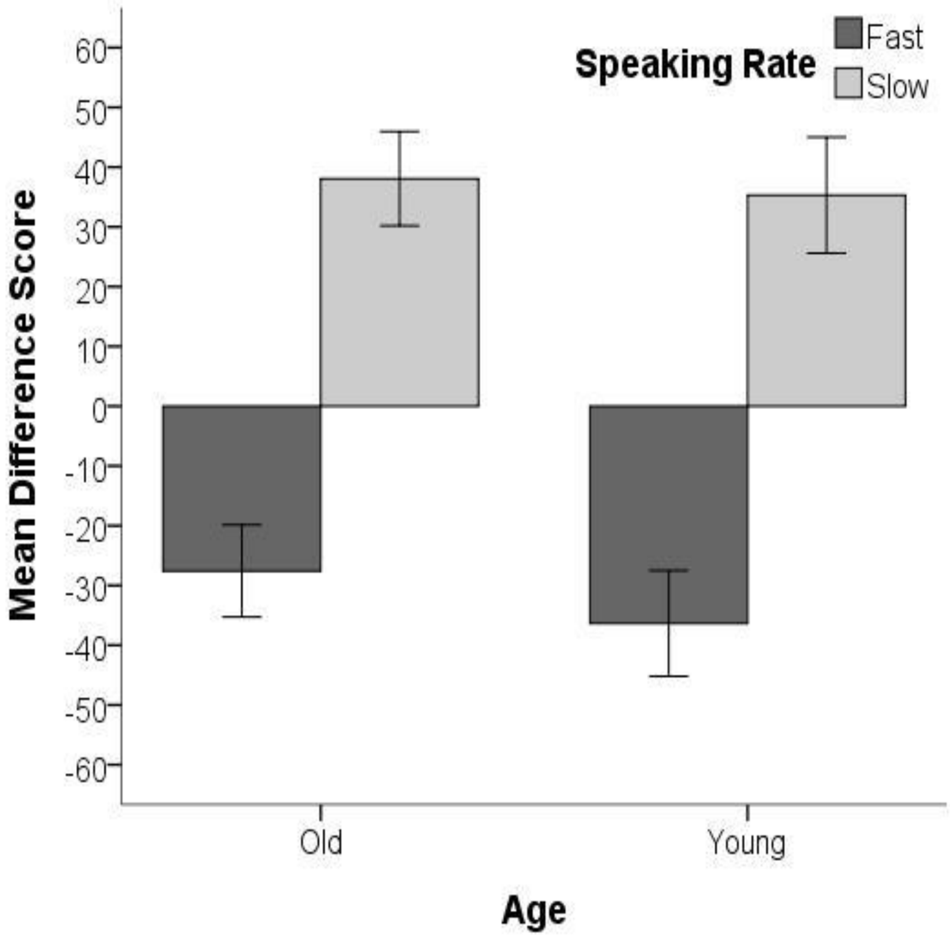
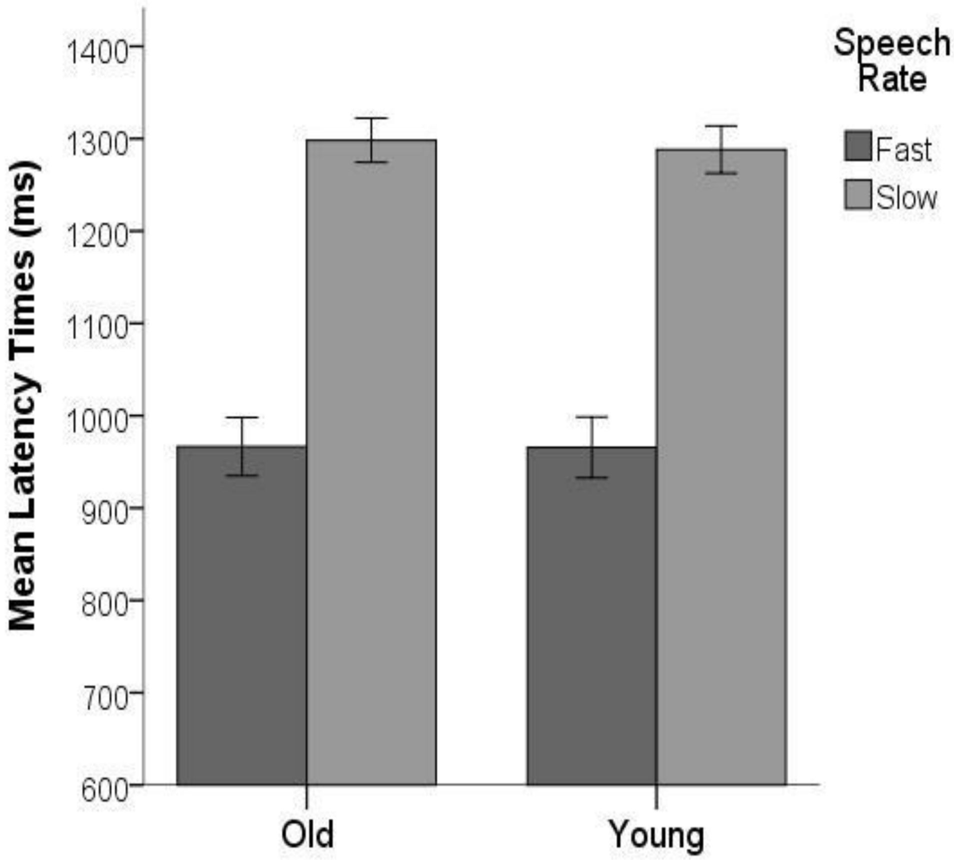

Speaking Rate
Fast
Slow
Mean Difference Score
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
Old
Young
Age

Figure 6



Mean Latency Times (ms)
1400
1300
1200
1100
1000
900
800
700
600
Old
Young
Speech Rate
Fast
Slow

Figure 7

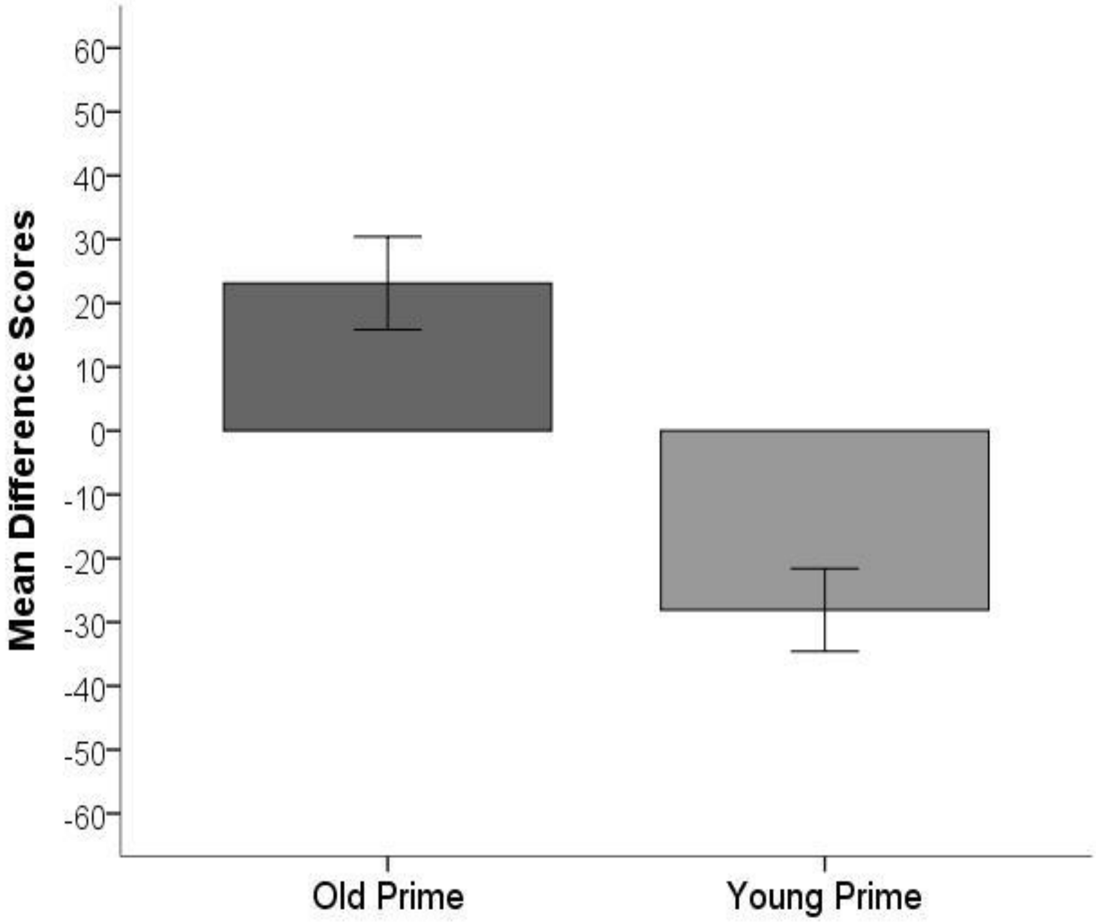
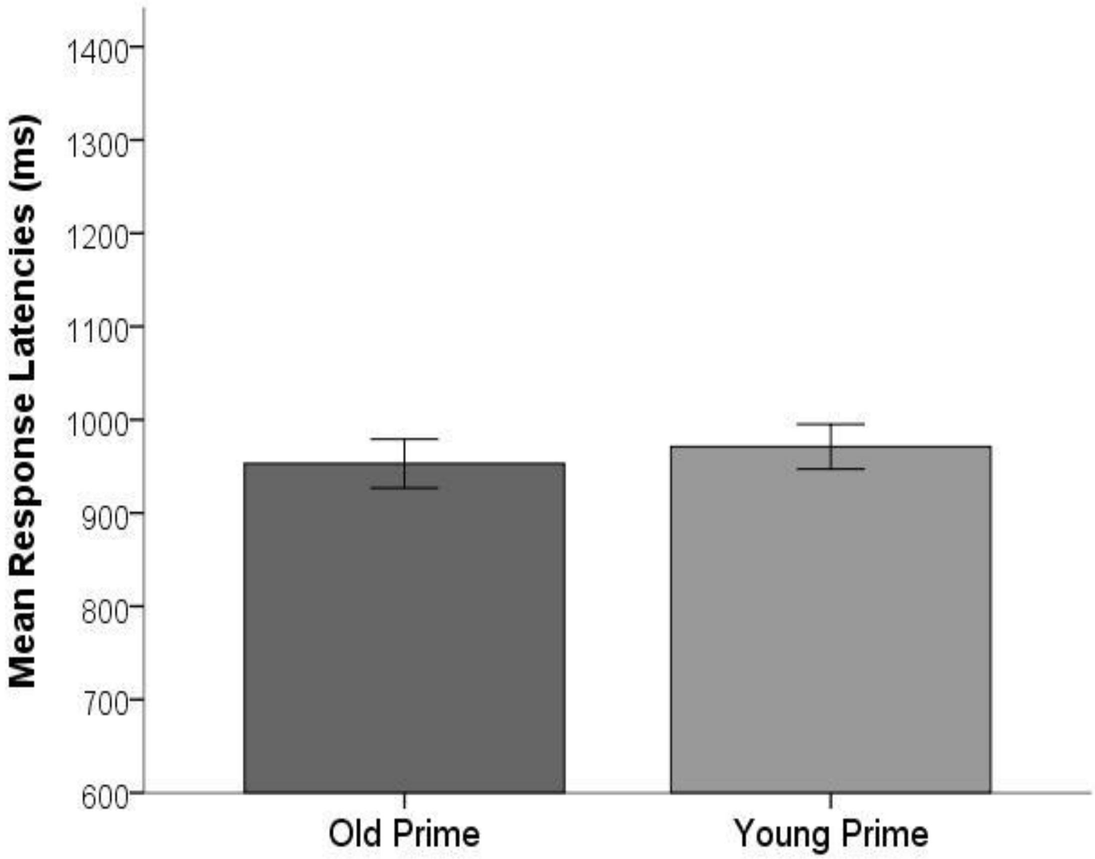

Mean Difference Scores
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
Old Prime
Young Prime

Figure 8


Mean Response Latencies (ms)
1400
1300
1200
1100
1000
900
800
700
600
Old Prime
Young Prime