



1

Introduction

1.1 Why Should We Care About Voice Quality?

Whenever we speak, our voices convey information about us as individuals. Speakers may sound young, or tired, or elated, or distracted. They may sound as if they are drunk, or lying, or ill, or bearing secret, exciting news. By their voices, adult speakers usually reveal whether they are male or female, and in addition, they may signal that they come from Texas, or Wisconsin, or France. Over the telephone or radio we may recognize the speaker as someone we know, or we may form a distinct impression of the physical appearance of someone we have never seen. The impressions listeners gain from voices are not necessarily accurate; for example, everyone has known the surprise of meeting a telephone acquaintance who does not match the mental picture we have previously formed of them. Despite such occasional mismatches, voice quality is one of the primary means by which speakers project their identity – their “physical, psychological, and social characteristics” (Laver, 1980, p. 2) or their “auditory face” (Belin, Fecteau, and Bedard, 2004) – to the world.

Table 1.1 non-exhaustively summarizes some of the kinds of judgments that listeners make when listening to voices. These human abilities arise from a long evolutionary process, and many animal species, including primates (Cheney and Seyfarth, 1980), wolves (Goldman, Phillips, and Fentress, 1995), penguins (Jouventin and Aubin, 2002), frogs (Bee, 2004), and bats (Balcombe and McCracken, 1992) use vocal quality to signal or perceive size, threat, and kin relationships. Human infants’ ability to recognize their mothers’ voices is in place at birth (DeCasper and Fifer, 1980), and responses to maternal voices can be measured *in utero*, suggesting such abilities develop even before birth (Hepper, Scott, and Shahidullah, 1993; Kisilevsky *et al.*, 2003). Voice conveys much of the emotion and attitude communicated by speech (Williams and Stevens, 1972; Banse and Scherer, 1996; Ellgring and Scherer, 1996; Van Lancker and Pachana, 1998; Breitenstein, Van Lancker, and Daum, 2001). Alterations in voice quality relative to the speaker’s normal vocal delivery may signal

Foundations of Voice Studies: An Interdisciplinary Approach to Voice Production and Perception,
First Edition. Jody Kreiman and Diana Sidtis.

© 2011 Jody Kreiman and Diana Sidtis. Published 2011 by Blackwell Publishing Ltd.



Table 1.1 Some kinds of judgments listeners make from voice.*Spoken message***Physical characteristics of the speaker**

Age
 Appearance (height, weight, attractiveness)
 Dental/oral/nasal status
 Health status, fatigue
 Identity
 Intoxication
 Race, ethnicity
 Sex
 Sexual orientation
 Smoker/non-smoker

Psychological characteristics of the speaker

Arousal (relaxed, hurried)
 Competence
 Emotional status/mood
 Intelligence
 Personality
 Psychiatric status
 Stress
 Truthfulness

Social characteristics of the speaker

Education
 Occupation
 Regional origin
 Role in conversational setting
 Social status

irony or sarcasm (Van Lancker, Canter, and Terbeek, 1981). Changes in rate and fundamental frequency affect the perceived “competence” (Brown, Strong, and Rencher, 1974) or credibility (Geiselman and Bellezza, 1977) of a speaker. Voice quality provides cues that indicate order of turn-taking in conversation (Schegloff, 1998; Wells and Macfarlane, 1998) and helps resolve sentential ambiguities (Kjelgaard, Titone, and Wingfield, 1999; Schafer, Speer, Warren, and White, 2000). Listeners may also judge the speaker’s sexual preference (Linville, 1998; Munson and Babel, 2007), status as native or nonnative speaker (Piske, MacKay, and Flege, 2001), and a myriad of personality factors (Scherer, 1979) based on voice quality cues.

This book describes the manner in which these kinds of information are conveyed to listeners, and how listeners draw conclusions – correctly or incorrectly – about speakers from their voices. Many of the points described are illustrated by recorded examples provided on the accompanying web site.

For example, consider the voice in audio sample 1.1. As you listen to this brief speech sample, you will probably automatically gather information about the speaker. Listeners agree that the speaker is female. Although opinions differ, listeners are likely to think that the speaker is adult but not elderly, cheerful, confident,

alert, and in good health. She is American, but does not have a pronounced regional, social, or ethnic accent. She sounds average or slightly above average in height and weight. She seems educated and is speaking carefully. She does not sound like a smoker. You probably do not recognize the voice, but it may remind you of someone you know.

Compare this talker to the voice in audio sample 1.2. This speaker is also female, but the voice sounds like a much older person. She has a New England accent, and the rhythm of her speech is unusual, making her sound rather upper-class or snobby (or merely self-conscious) to some listeners. She is not tired, depressed, or angry, but she is not obviously happy, either, and may be bored. Her voice is somewhat hoarse, suggesting that she is or has been a smoker, but she does not seem ill. Listeners disagree somewhat about her height and weight, but generally estimate that she is average or slightly below average in height, and slightly above average in weight.

The voice of a speaker with a vocal pathology is presented in audio sample 1.3. Even this short sample may produce complex impressions of old age, illness, and unattractiveness, along with a sense of the speaker's emotions or mood, intelligence, and competence. Patients who develop a voice disorder often complain that the disordered voice is not really their voice, and does not convey who they are. In some cases, patients dislike the image they portray so much that they avoid speaking, resulting in significant social and work-related difficulties. Severe voice quality problems may also interfere with speech intelligibility, creating a handicap in the communication of verbal information (Kempler and Van Lancker, 2002).

The strong impressions conveyed by voice quality are often manipulated by the media for multiple purposes. For example, in the classic film *Singin' in the Rain* (Freed, Kelly, and Donen, 1952), the shrill, loud voice of the character Lina Lamont (played by actress Jean Hagen) surprises and amuses because it does not fit her appearance (a beautiful, smiling blonde) or the elegant, poised, sophisticated personality she visually projects. This contrast – a prototypically silly voice in an elegant physique – forms a running joke throughout the film, playing off such lines as, “What do you think I am, dumb or something?” spoken in the abrasive voice stereotypically associated with a vulgar, uneducated, shrewish female. More often, voices are selected to fit the intended message. Documentary films enhance credibility through the use of a male narrator whose voice carries the stereotype of an authoritative figure who is solid, mature, calm, highly intelligent, and dignified. In the field of advertising, impressions conveyed by voice quality are integral to establishing a product image. Consider the characteristics projected by the voices typically used in advertisements for luxury automobiles. Low pitch, breathy quality, and a fairly rapid speaking rate produce the image of an intimate message from a mature but energetic male who possesses authority, sex appeal, social status, and “coolness.” These vocal attributes are appropriate to the economic niche for the product and imply that its owners are powerful, sexy, and affluent.

Given the wide range of information listeners derive from voices, it is not surprising that scholars from many different disciplines have studied the production and perception of voice. Table 1.2 lists some of these disciplines, along with a sampling of typical research questions. These research questions encompass much of human existence, and indicate how central voice quality is to human life.

Table 1.2 Disciplines incorporating the study of voice and voice quality.

<i>Discipline</i>	<i>Some typical research questions</i>
Acoustics	Deriving reliable and meaningful acoustic measures of voices
Animal behavior	Vocal recognition of kin and social information by nonhuman animals
Biology	Biological and evolutionary significance of vocalization
Computer science, signal processing, information	Transmission, measurement, and synthesis of voice
Forensic science, law enforcement	Reliability and verification of "earwitness" testimony; assessment of truthfulness from voice
Linguistics, phonetics	Meanings of vocal quality in speech
Medicine:	
Developmental biology	Infant voice recognition
Gerontology	Voice quality changes in aging
Neurology	Brain function underlying vocal behaviors
Obstetrics	Prenatal voice perception
Otolaryngology	Voice disorders
Pediatrics	Childrens' processing of vocal information
Physiology	Control of phonation
Respiration	Role of breathing in vocalization
Surgery	Effects of surgical interventions in the vocal tract on voice; cosmetic changes for transgendered voices
Music:	
Singing	The singing voice: many questions
Vocal coaching	The effects of training on the voice
Physics	Vibrating laryngeal tissues; relation of vibration to sound; patterns of airflow through the glottis
Psychology:	
Cognitive psychology	Speaker recognition and its causes; interaction between speech recognition and voice quality
Clinical psychology	Detecting depression, psychopathology, and personality in the human voice
Social psychology	Voices as signals of social relationships including conversational turn taking, sarcasm, and successful con-artistry
Neuropsychology	Brain mechanisms underlying the perception and production of voice cuing personal identity as well as mood and motivation
Psychophysics	Relevant acoustic voice features for perception
Psycholinguistics	Voice information in meaning comprehension for grammatical structure and nonliteral meanings
Sociology	Voice types associated with social groups and their development
Speech science	Normal voice and speech production
Speech pathology	Effects of vocal pathologies on voice quality
Theater arts	Voice as artistic instrument



3

Neurological Foundations of Voice Production and Perception

3.1 Introduction

Our goal in this chapter is to assemble fundamental information on cerebral processes underlying various aspects of vocal production and auditory perception that will serve as neurological foundations to voice studies. More detail on these topics is presented in Chapter 6. We begin with essentials of vocal production.



3.1.1 Our personal voice pattern: Producing a voice takes a whole person



Personal characteristics of broad palette are carried in the voice. Any given vocal performance requires a complex orchestration of the vocal apparatus, deriving from expansively represented cerebral influences. Subcortical (including limbic) systems of the brain generate the attitudes, moods, and emotions that find expression in our voices. Temporal and parietal lobes accumulate memories and facts leading to beliefs, assumptions and opinions, which also influence pitch contours and accent, speaking rate, loudness, and other vocal attributes. Intentions, goals and motor programs are managed by frontal lobe systems in circuitry with the subcortical nuclei that contribute to every vocal effort. It is a remarkable fact that this cognitive cornucopia emerging from whole brain interactions is expressed in the human voice. It is not known in detail how the limbic system (seat of emotions), subcortical nuclei (motoric and motivational control) and the cortical lobes (engaged in cognition and planning) operate convergently to formulate the psychological characteristics that are manifest as attitude, mood, and personality, but it has been repeatedly demonstrated that all of these are “heard” along with personal identity in the voice, leading to the oft-repeated statement in medical (Woodson, 1999),

Foundations of Voice Studies: An Interdisciplinary Approach to Voice Production and Perception,
First Edition. Jody Kreiman and Diana Sidtis.
© 2011 Jody Kreiman and Diana Sidtis. Published 2011 by Blackwell Publishing Ltd.



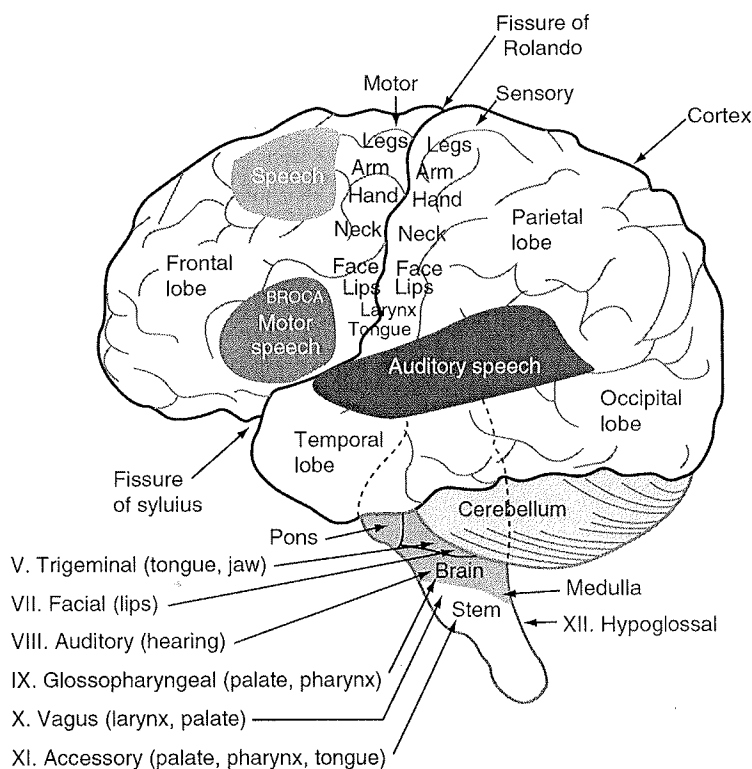


Figure 3.1 Schema of brain structures involved in vocalization. Important structures not seen are the basal ganglia and limbic structures below the cortical mantle, and the periaqueductal grey matter in the midbrain.

religious¹ (Cumming, 1856), and literary writing (Longfellow, 1866) that voice is revelatory of mind or soul. The voice is also a product of our evolutionary past (Locke, 2009). In this chapter we also touch on the controversies surrounding whether human vocal behaviors are significantly related to those of nonhuman animals, and if so, in what ways.

3.1.2 Overview of brain structures and functions in vocal production: It takes a whole brain

As our introductory remarks suggest, producing a vocal pattern requires contributions from nearly the entire brain (see Figure 3.1). Vocalization, which requires “remarkable coordination,” draws on functionality across most of this neural resource (Dubner, Sessle, and Storey, 1978, p. 211; Jürgens, 2002). Respiratory,

¹ “There is something in the human voice, just as in the human face, that I have often thought is one of the greatest proofs of the infinitude of the resources of God” (Cumming, 1856, p. 80).