UNCOVERING THE ROLE OF GENDER STEREOTYPES IN SPEECH PERCEPTION

ELIZABETH A. STRAND
The Ohio State University

This work examines the effect of gender stereotypes on the perception of language by drawing together findings from the fields of speech perception, gender studies, and social psychology. Results from two speech perception experiments are reviewed that show that listeners' stereotypes about gender, as activated by the faces and voices of speakers, alter the listeners' perception of the fricatives /s/ and /ʃ/ . One experiment employs auditory-only consonant-vowel-consonant (CVC) tokens and the other employs audiovisual stimuli created from the same tokens synthesized with talking faces. This effect of stereotypes on low-level speech processing must be accounted for in models of perception, cognition, and the relationship between the physical and social environment.

Much work on the sociology of language has demonstrated conclusively that perceived language variation triggers evaluative judgments about the speaker. Furthermore, the independence of these judgments from the percept of language variation is usually taken for granted. This article, however, discusses how evaluative judgments, or stereotypes, actually play a role in shaping the perception of language itself. As von Hippel, Sekaquaptewa, and Vargas (1995) note,

> Any evidence that perceptual processes influence and are influenced by stereotypes and prejudice would have profound implications. People view their senses as documentary devices that faithfully translate the environment into understandable and manageable units . . . they *accept* what they see and hear. (p. 181)

AUTHOR'S NOTE: *This article is based on work supported by the National Institute on Deafness and Other Communication Disorders under Grant No. 7R29DC01645-04, awarded to Keith Johnson, as well as a Cognitive Science 1997 Summer Fellowship Grant awarded to the author jointly by the Center for Cognitive Science, the Department of Linguistics, and the Department of Spanish & Portuguese at the Ohio State University. Much gratitude is owed to those who provided input on various aspects of this work, including Mary Beckman, Julie Boland, Penny Eckert, Keith Johnson, Norma Mendoza-Denton, Frederick Parkinson, Bill von Hippel, and Don Winford. Thanks also to Dennis Preston and Lesley Milroy for their assistance. Requests for reprints should be sent to Elizabeth A. Strand, the Ohio State University, Department of Linguistics, 1712 Neil Avenue, Columbus, OH, 43210-1298.*

JOURNAL OF LANGUAGE AND SOCIAL PSYCHOLOGY, Vol. 18 No. 1, March 1999 86-99

A half century of work on speech perception has clearly demonstrated the richness of the phonetic information that listeners use in processing spoken language. In addition to stop burst spectra and consonant-vowel formant transitions, for example, listeners also use visual information about the place of articulation of consonants in perceiving speech (e.g., McGurk & MacDonald, 1976). Strand and Johnson (1996), however, showed that beyond just visual phonetic or articulatory information about the segments produced, information about the gender of a talker (e.g., spectral characteristics of the voice and characteristics of the face) also influences speech perception. That is, expectations of gender prototypicality and stereotypes can change listeners' perception of the physical signal.

As Strand (1997) points out, this implies that socially constructed beliefs or expectations (i.e., stereotypes) about how a speaker "should" sound, based on how he or she appears, are actively engaged in speech perception. That is, bottom-up processing of acoustic information directly interacts with higher level information related to people's socially constructed stereotypes about gender. If it is true that stereotypes affect even our basic categorization of the speech signal, then a whole class of cognitive theories that maintain that bottom-up sensory processes and top-down cognitive processes do not directly interact with each other would become untenable.

SPEECH PERCEPTION AND SPEAKER NORMALIZATION

PERCEPTUAL NORMALIZATION

As Peterson and Barney (1952) note, there is a lack of acoustic invariance in the speech signal as produced by different talkers. The same word or syllable as uttered, for example, by a female and male talker will have different acoustic characteristics due to physiological differences, as well as aspects of voice quality that are salient for an individual's gender identity within a specific speech community.

Vowel formant frequencies will in general be lower, bandwidths will be wider, and fundamental frequency will be lower for a male talker. Such intersex acoustic differences are due to the generally larger size of the adult human male vocal tract when compared with that of the adult human female, and the descent and reshaping of the male larynx at puberty, as well as to culture-specific conventions about habitual postures of the voice and vocal tract for different genders (see, e.g., Graddol & Swann, 1989, chap. 2; Mendoza-Denton & Strand, 1998; van Bezooijen, 1995).

Considerable research has shown, however, that listeners compensate for differences among speakers that result in varied acoustic

outputs to "hear" tokens of the same word or syllable as the same linguistic category (e.g., Bladon, Henton, & Pickering, 1984; Johnson, 1991; Ladefoged & Broadbent, 1957). Any theory of speech perception must account for this *perceptual normalization*.

The gender-related variability targeted in this article, namely, the distribution of energy in the spectra of fricatives /s/ and /ʃ/, nicely illustrates the complexity of the relationship between physiology and convention. The /s/ spectrum is concentrated at a higher frequency than the /ʃ/ spectrum, reflecting the smaller size of the resonating cavity in front of the sound-generating constriction for /s/. This phonetic contrast rides on top of the gender-related variation: Men generally have lower frequency turbulence than women.

In considering the differences between female and male production of the fricatives /s/ and /ʃ/, which is important for the results of Strand and Johnson (1996), the variation in spectra across sexes is often attributed to the relative length or size of the male versus female vocal tract (e.g., Schwartz, 1968, who refers to the vocal tracts of men as "bigger" and those of women as "smaller").

All models of fricative consonants have assumed that "[fricative] sound [is] predicted to be generated in and downstream of the constriction" (Shadle, 1991, p. 412). But general female-male vocal tract size differences, however, exist mainly behind the area of the constriction and the obstacle (upper teeth for /ʃ/, lower teeth for /s/) necessary to produce the fricatives, and the lips (sound "radiators").

The existence of such minimal physiological differences in front of the constriction in females' and males' fricative productions raises the question of what, then, is affecting such fricative differentiation between general sex categories. In her proposal of a revised model of fricative production, Shadle (1991) expands the traditional understanding of fricative acoustics by stating that "differences too small to affect the area function significantly, that had no effect on the constriction, flow rate, or pressure drop, nevertheless could have a substantial effect on the radiated sound . . . minute abrupt discontinuities [cause] significant acoustic differences" (p. 423).

In a descriptive study of the differences between American English-speaking female-male production of /s/, Naslund (1993) found that in general, the women he studied tended to use a more fronted, slit variant of /s/, whereas the men he studied tended to use a more alveolar, grooved variant. In addition, Naslund found that by age 8, a large majority of boys and girls seem to produce the fricative in a manner consistent with the adult production of the variant linked to their gender group. This result is particularly interesting in view of the similarity of vocal anatomy of prepubescent girls and boys.

Such evidence points to the development and salience of socially influenced fricative productions, at least in the American English-speaking groups studied here. The fact that children show acquisition

of these specific gendered pronunciations at such a young age coincides with other research that indicates that children acquire salient gender-linked linguistic traits very early in the process of language acquisition (see, e.g., Coates, 1993, chap. 7).

However, despite such interspeaker acoustic variation even in very young speakers, listeners still perceive speech accurately. Listeners normalize speech through reference to experience-based expectations regarding speaker-to-speaker variation.

The notion of experience-based expectation was first suggested by May (1976), who posits a physiologically conditioned expectation to justify the location of the perceived boundary between /ʃ/ and /s/ on a fricative continuum. May (1976) presented listeners with fricative-vowel tokens that were constructed from synthetic fricatives on a 10-step continuum between /ʃ/ and /s/, spliced with a following naturally produced vowel. Listeners were asked to identify tokens as either /ʃ/- or /s/-initial. Results indicate that the boundary between the two fricative categories shifts in frequency depending on whether a male voice (having the properties of a larger vocal tract) or a female voice (having the properties of a smaller vocal tract) produced the following vowel portion in the token. May concluded that the lack of invariance in the cues for the categorization of /s/ and /ʃ/ is accounted for by listeners through normalization for vocal tract size based on the characteristics of the following vowel.

So essentially, listeners had one expectation for a female talker regarding the frequency location of a boundary between the categories of /ʃ/ and /s/, and a different expectation for a male talker. This suggests that perceptual normalization must be viewed as a sophisticated dynamic process that can accommodate varying percepts of individual talkers.

UNCOVERING A GRADIENT SPEAKER NORMALIZATION EFFECT

The research discussed above implies that the gender of talkers will be perceived to be in a dichotomous relationship; Strand and Johnson (1996), however, found evidence to suggest that the sex of a voice is perceived gradiently. In one experiment of this study, results from a fricative categorization task suggest that even *within* sex categories, voices can affect normalization processes in significantly different ways.[1]

Prior to constructing auditory stimuli for this experiment, naïve listeners judged 37 voices on a unidimensional measure of gender prototypicality. Four voices were selected for stimuli construction: one judged as most prototypically female sounding; one judged as most prototypically male sounding; one judged as most nonprototypically female sounding; and a final voice judged as most nonprototypically male sounding.

Listeners accurately judged the sex of each voice in each case, with none of the nonprototypically judged voices being mistaken for sex, as prior research has indicated may occur (see, e.g., Edwards, 1979; Lass, Almerino, Jordan, & Walsh, 1980).

A nine-step fricative continuum from /ʃ/ to /s/ was then synthesized. These nine fricative noises were concatenated with naturally occurring vowel-consonant (VC) coda chunks from the words *sod* and *shod* as spoken by the four speakers identified above. The synthesized stimuli were accepted by naïve listeners as sounding like naturally produced speech. Figure 1 provides a schematic spectrogram of these tokens.

Listeners were presented with the tokens and asked to identify each stimulus as either the word *sod* or the word *shod*. Results of this voice-only experiment indicate that talker normalization of the fricatives is gradient with respect to the listeners' (auditory) perception of the gender of the speaker. Figure 2 presents these results, showing averaged identification functions for the four voices. The four category boundaries are statistically reliably different from each other (the 50% point on the vertical axis represents the category boundary between /ʃ/ and /s/)—listeners appear to be normalizing the synthetic fricatives differently depending on the voice producing the rest of the word. Reflective of the link described above between the points made in Schwartz (1968) and May (1976), longer vocal tracts (i.e., those typical of men) induce boundaries at lower frequencies, and shorter vocal tracts (i.e., those typical of women) induce boundaries at higher frequencies.

However, speakers are not simply "normalizing" these fricatives to one or the other gender category. Rather, they are employing more finely tuned expectations within gender categories, leading to boundary locations even within the categories. These results indicate that listeners likely employ a good deal of "indexical" information from the speech signal in normalization and perception, information that has traditionally been considered much more relevant for signaling speaker identity than for use in phonological categorization.

This experiment found that voices perceived as sounding more or less prototypical with respect to gender attributes produce stronger or weaker (respectively) perceptual normalization effects.[2]

UNCOVERING A CROSS-MODALITY NORMALIZATION EFFECT

While the McGurk Effect (McGurk & MacDonald, 1976) shows that visual and auditory phonetic cues are perceptually integrated, the work in Strand and Johnson (1996) examines whether indexical information, notably expectations about gender activated by talkers' faces, similarly affects perception.

Strand and Johnson's (1996) second experiment examined the role of visual information about talker identity in perceptual normalization

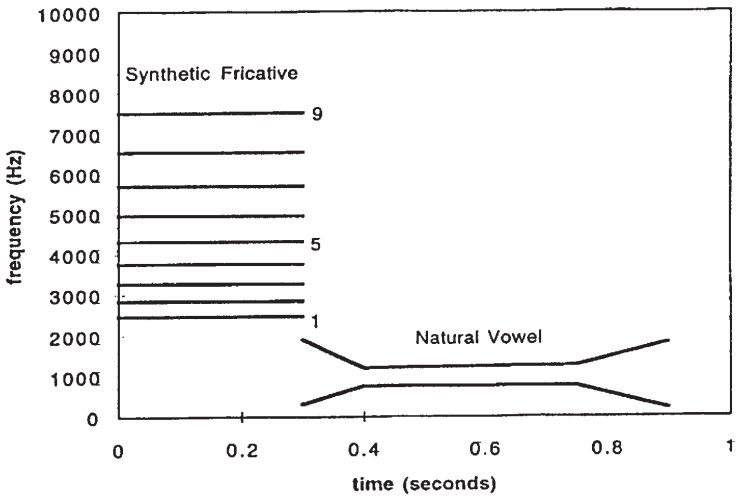

Figure 1. **Schematic spectrograms of the tokens 1-9, [ʃad] ↔ [sad].**
Note. Synthetic fricative steps are indicated in equal Bark units (an auditory scale related logarithmically to the more familiar acoustic Hz scale); the fricative measurement indicated is the fricative pole value. The F1 and F2 patterns of the vowel portion of the token are schematized in the "natural vowel" portion of the token representation.

by pairing the audio tokens of the first experiment with different male and female faces, to make audiovisual tokens that were short movies of faces saying the two words.

The method employed in this experiment included first digitizing movies of prototypical male and female faces saying *sod*. The token continua for the nonprototypical male and female voices from the first experiment were then dubbed onto these movies, both crossing and matching faces and voices for gender. So half of the audiovisual tokens had face and voice sexes that were matched, and half had sexes that were crossed.

A second set of participants were presented with the audiovisual tokens, again in a two-alternative forced-choice paradigm. Results indicate that there is indeed integration of the visual information about gender of the speaker (extracted from the face) with the acoustic information in the speech signal. That is, the same fricative was perceived differently depending on whether it was accompanied by a male or a female face producing the token. Similar to what was found in the auditory-only experiment, a female face induced the perceived boundary between /ʃ/ and /s/ to shift up in frequency, whereas a male face induced the boundary to shift down in frequency. This effect will be refered to as the "Face Gender Effect."

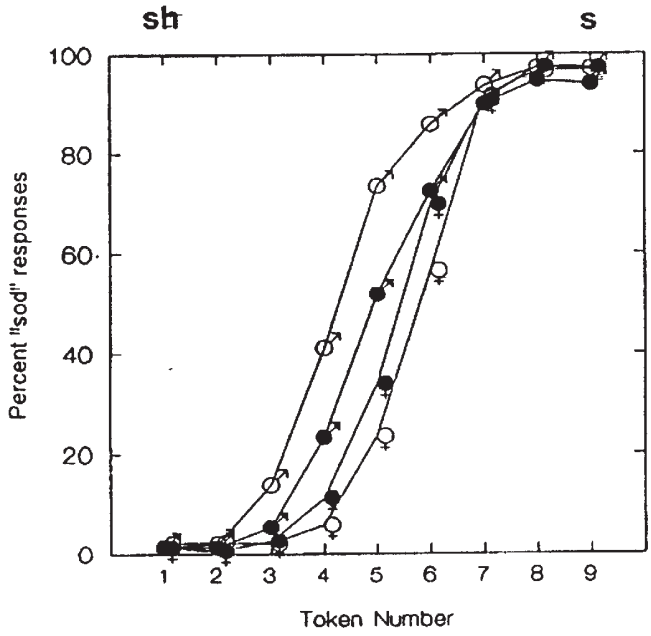

Figure 2. **Results of the voice-only experiment: A gradient talker normalization effect can be seen.**
Source. From Strand and Johnson (1996, p. 19).
Note. Responses to the prototypical male and female talker are plotted with open Mars and Venus signs, respectively, and responses to the nonprototypical talkers are plotted with filled symbols. The x-axis represents the fricative Tokens 1 through 9, and the y-axis represents the percentage "sod" responses to the tokens.

From these audiovisual results, it can be concluded that talker normalization definitely involves the integration of auditory and visual information, similar to the McGurk Effect, but more so, it involves accessing the expectations (stereotypes) that we have about how people should sound, based on how they look.

The Face Gender Effect has also been replicated using audiovisual tokens and a vowel continuum spanning the vowels in *hood* [hʊd] to *HUD* [hʌd] in which F1 values were altered for each step in the continuum (Johnson & Strand, 1998). Pariticipants tended to perceive the boundary between vowel categories at a lower F1 frequency for tokens that were concatenated with a male face and at a higher F1 frequency for tokens concatenated with a female face. These results are consistent with listeners' likely expectations that, for example, a specific F1 value that is appropriate for a token of *hood* as produced by a female speaker would be too high for a token of the same word as produced by a male speaker, for whom it would be perceived as an example of *HUD*.

These results indicate that higher level, relatively complex social expectations might have an influence on such low-level basic processes as phonological categorization of the speech signal. This finding challenges "gender-free" conceptions of human speech perception and language processing, and introduces a new set of complications to the present conception of cognitive processing.

RELEVANT PARAMETERS FOR A THEORY OF GENDER STEREOTYPES & SPEECH PERCEPTION

> Although some of the assumptions made [within] the perceptual system are inborn, others are learned through experience. . . . Consequently, because our knowledge of the world guides what we see and hear, stereotypes have a great deal of potential to influence perception. (von Hippel et al., 1995, p. 181)

The issues discussed in the previous sections of this article indicate the necessity for a new theory on which to base and test hypotheses about the effect of gender information on speech processing. Current speech perception theories that do not allow for the influence of information, such as that which the face of the talker may provide, cannot offer a fully inclusive account of the perception process. Developing such a theory necessarily involves approaches and data from a multitude of research areas including knowledge about socially constructed, culture-specific gender identity and resulting stereotypes; the way in which facial features have been shown to activate certain gender expectations; the cognitive processes involved in accessing and using stereotype information in information processing; and finally, the incorporation of gender stereotype information into speech recognition.

STEREOTYPE RESEARCH

Research on stereotype formation, encoding, and access falls within the realm of social psychology and its subfield of study, person perception. The study of person perception includes inquiry into how people make impressions of others, what information biases these impressions, what kinds of information are used in forming impressions, and how accurate these impressions actually are (Taylor, Peplau, & Sears, 1994, p. 35).

Before stereotypes are formed, perceivers construct general impressions of people and groups of people on the basis of a number of different types of information, such as physical cues about gender and race, observable behavior, and judgments of competence and sociability. People tend to infer personality traits from cues such as these in a very automatic way. As Taylor et al. (1994) point out, the inferences that

people form of others based on minimal information are easily integrated to form overall impressions of people that remain active for long periods of time (pp. 35-36).

The cognitive approach to person perception and impression formation highlights the notion that the abilities of people to deal with incoming sensory stimulation are limited, so perceivers therefore absorb information only selectively and then organize it into a meaningful "gestalt" that helps them to make sense out of the entire input (Taylor et al., 1994, p. 42). One efficient mechanism for organizing incoming information is to categorize each new stimulus as soon as possible, to give a shortcut into perceivers' worldview.

From basic impressions of others come stereotypes, which are generally considered to be beliefs about the characteristics of a certain group. Beyond just beliefs about personal attributes, however, von Hippel et al. (1995) point out that stereotypes are also theories "about how members of another group look, think, and act, and how and why these attributes are linked together" (p. 178). A stereotype

> guides behavior and judgment . . . [and] may be represented in a variety of fashions: as a social category, a schema, a base rate, a distribution, an expectancy, a prototype, an exemplar, an associative network, or even perhaps a vast collection of instances. (p. 178)

As this last phrase indicates, social psychologists vary in how they attempt to model stereotypes, but all would agree that stereotypes have a strong influence in perceiving social information. The research reviewed in the previous section strongly suggests that stereotypes also play a role in perceiving speech.

It is important to note, as Fiske and Taylor (1984, p. 165) point out, that "categorizing others leads to exaggerating perceived differences between groups and minimalizing perceived differences within each group" (p. 165). So, even though stereotypes about groups of people are formed and maintained to help perceivers understand reality, these stereotypes can often induce false perceptions of reality.

THE PSYCHOLOGY OF GENDER AND GENDER STEREOTYPES

Research in the psychology of gender and gender stereotypes that exist within specific cultures will also have bearing on building a theory of perceived gender influences on speech processing. The categorization of people according to their sex is one of the most primary judgments that human perceivers make of each other. Berscheid (1993) describes the potency of gender information to structure human behavior by stating that "knowledge of another's gender identity is the master key that immediately unlocks a . . . vast array of beliefs and

stereotypes about that person's nature" (p. ix). Accordingly, we react to people as we see fit, based on these expectations.

Geis (1993) continues the argument that sex and gender stereotypes influence our perceptions of others. She states that stereotypes

> enhance perceptions, interpretations, and memories that are consistent with stereotypical attributes and obscure, diffuse, or cause us to disregard or forget information that is inconsistent with them. . . . Thus, even when women and men behave alike, we see them as different. (p. 12)

These notions, coupled with the assertion made by von Hippel et al. (1995) that stereotypes "change what is actually perceived in the world" (p. 180), give powerful support to the hypothesis that social expectations about gender have an effect on auditory processing of speech.

FACE PERCEPTION RESEARCH

Social psychologists are also interested in research on the perception of faces within and across cultures. As Zebrowitz and Montepare (1992) state, much research

> indicates that appearance has a significant impact on social perceptions and interactions . . . not only is appearance one of the initial qualities that children and adults mention in their descriptions of people, but it also influences the traits they ascribe to others. (p. 1143)

Face perception researchers have been able to successfully correlate specific facial features and orientations of groups of facial features with specific personality attributes ascribed to the people whose faces perceivers are presented with. In addition, within a culture, perceivers seem to be consistent in the personality attributes that they ascribe in response to specific facial features.

For example, Berry and Zebrowitz McArthur (1985) show how physical appearance variables influence stereotypical impressions of character attributes in a social-psychological effect called the *Babyface Effect*. These researchers first made careful measurements of specific facial features and were able to correlate those features and combinations of features with perceptions of certain personality characteristics. They found that certain factors such as overall low placement of facial features, large eyes, and a short nose increased the perceivers' perceptions of, for example, physical weakness, social submissiveness, naïveté, warmth, and honesty. These personality characteristics have long been considered infantile psychological attributes, as well as stereotypically female. So they concluded that people who exhibit "babyface-like" facial features such as these tend to thereby be ascribed with such psychological attributes.

In terms of activating stereotypes of various personality characteristics, social psychologists have also investigated issues such as notions of what constitutes facial attractiveness within a culture, the effect of facial attractiveness on perceived typicality, and the effect of race and age on activation of certain stereotypes.

IMPLICATIONS FOR CURRENT THEORIES

Results such as the Face Gender Effect have general implications for cognitive theories at all levels of processing, from the most specific (e.g., theories about speech signal processing) to the most general (e.g., theories about the general ordering and interaction of cognitive processes, such as modularity theory). Such theories must be reexamined in terms of their ability to account for these effects.

LINGUISTICS

In terms of our basic linguistic assumptions about phonological categorization, that is, that such categorization is sensitive only to relevant phonetic information, the Face Gender Effect is problematic. This study demonstrates that socially constructed beliefs do play a role in even the basic phonological categorization of both vowels and consonants. In the most broad sense, these results are incompatible with an impoverished view of the phonetics-phonology interface such as that set forth in Chomksy and Halle (1968), in which language-specific phenomena reside within a minimalist categorical phonological component that interacts with "universal phonetics."

Although more recent notions of linguistic systems may support language-specific phonetics, the studies and results discussed here indicate that any component that categorizes the speech signal must be far richer than was originally conceived three decades ago. In suggesting that socially constructed information also affects how we perceive and categorize speech sounds, these results force us to consider an even more complex interaction of salient information than has been done in the past.

LANGUAGE AND GENDER, VARIATIONIST SOCIOLINGUISTICS

Similarly, the notion that gender stereotype information influences variability in the very organization of the perceptual space also has an effect on traditional approaches to sociolinguistic variation and sound change, as well as language and gender study. Although sociolinguists have embraced the notion of the heterogeneity of grammar across

speakers, the indication that perceptions of gender both influence and constrain the processing of low-level sensory input, as in the Face Gender Effect, forces us to rethink the notion of what the categories of input actually are in the first place and what information they include or at least have access to.

As the concept of lexical diffusion was groundbreaking in implying heterogeneity in the lexicon in terms of driving sound change, a phenomenon such as the Face Gender Effect goes even further to imply heterogeneity within a single category of input into the grammar. Essentially, this implies that our basic sound categories will have to vary depending on whom we perceive to be producing those speech sounds. So the question of what phonological categories actually are is again raised. The ways in which this issue is addressed are important for consideration within sociolinguistic theory.

SPEECH PERCEPTION

Theories of speech perception, which have not traditionally dealt with listeners' social expectations, will also require modification to explain and predict the influence of socially constructed beliefs and stereotypes on perception. Although it could be argued that some views of normalization might be able to handle these results, such as the theory posited by Bladon et al. (1984) (in which the speaker is first identified as male or female with that identification then guiding a sliding auditory expectation scale), a closer examination of the predictions made by these theories indicates that they cannot account for effects such as the perceptual gradiency within voice gender categories that was originally reported in Strand and Johnson (1996) nor any influence of the Face Gender Effect.

In a related vein, theories of perceptual learning within a general information processing view of cognition are based on the notion that perceptual categories are learned, with the relationship between perception and categorization developed through a continual process of "learned associations." In the case of the learning of phonological speech perception categories, most of the focus to this point has been on the influence of acoustic information alone. These results, however, suggest that learning these categories involves a visual as well as an auditory component, and more specifically a visual component that goes well beyond simply extra-acoustic information about place of articulation such as that displayed in the McGurk Effect. The Face Gender Effect indicates that subtle connections are learned across, not just within, modalities. Therefore, the process of perceptual learning is much more complex than had originally been formulated.

COGNITION

Turning to general theories of cognition, modularity theory (Fodor, 1983) exhibits a basic inability to handle the Face Gender Effect results discussed here. As Garfield (1987) summarizes, modularity theory is based on the hypothesis that

> the mind is not a seamless, unitary whole whose functions merge continuously into one another; rather, it comprises—perhaps in addition to some relatively seamless, general-purpose structures—a number of distinct, specialized, structurally idiosyncratic modules that communicate with other cognitive structures in only very limited ways. (p. 1)

Fodor's theory supposes a number of constraints on the nature and allowable interaction of these modules, including the constraint of "information encapsulation," a primary tenet of modularity. Information encapsulization implies that input systems such as those for low-level speech processing are informationally isolated from the higher level central processes. The Face Gender Effect, however, suggests that information derived from the central processes (in this case, gender expectations or stereotypes) must directly affect processing within the input systems (here, low-level speech processing), thereby defying the strong view of modularity's claims, which says that these systems are mandatorily held separate.

Remez (1994) discusses results indicating that listeners' belief of a signal to be speech (rather than being simply a sine wave) affects whether they perceive it as speech, and makes a claim that holds in the present case as well: "All proponents of a modular view of speech perception hold to encapsulization, which [such] evidence falsifies by *delineating a role of belief* in the creation of a phonetic perceptual mode" (p. 161, emphasis added).

CONCLUSIONS

The results presented in Strand and Johnson (1996) suggest that any model of speech perception that is rigidly modular with a quasi-symbolic, discretely categorized output just above the level of sensation, with gender stereotypes then acting on this output, will not work. These results indicate that variability in the speech signal is not just normalized away but is actually used in speech processing. Therefore, representations of both speech categories as well as stereotypes about gender, nationality, race, and so forth must necessarily be interconnected. More traditional ways of dealing with interspeaker variability in speech perception will not work.

NOTES

1. The methods and results of Strand and Johnson (1996) are only briefly described here. Please refer to the original source for more explicit details.

2. These perceptual results support the arguments against the use of broad-category acoustic averages for the description of gender differences of voices, as laid out in Mendoza-Denton and Strand (1998).

REFERENCES

Berry, D. S., & Zebrowitz McArthur, L. (1985). Some components and consequences of a babyface. *Journal of Personality and Social Psychology*, *48*(2), 312-323.

Berscheid, E. (1993). Forward. In A. E. Beall & R. J. Sternberg (Eds.), *The psychology of gender* (pp. vii-xvii). New York: Guilford.

Bladon, R.A.W., Henton, C. G., & Pickering, J. B. (1984). Towards an auditory theory of speaker normalization. *Language & Communication*, *4*(1), 59-69.

Chomksy, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row.

Coates, J. (1993). The acquisition of gender-differentiated language. In *Women, men, and language* (chap. 7, pp. 141-167). London: Longman.

Edwards, J. R. (1979). Social class differences and the identification of sex in children's speech. *Journal of Child Language*, *6*, 121-127.

Fiske, S. T., & Taylor, S. E. (1984). *Social cognition*. New York: Random House.

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Garfield, J. (Ed.). (1987). *Modularity in knowledge representation and natural-language understanding*. Cambridge, MA: MIT Press.

Geis, F. L. (1993). Self-fulfilling prophecies: A social psychological view of gender. In A. E. Beall & R. J. Sternberg (Eds.), *The psychology of gender* (pp. 9-54). New York: Guilford Press.

Graddol, D., & Swann, J. (1989). The voice of authority. In *Gender voices* (chap. 2, pp. 12-40). Cambridge, MA: Blackwell.

Johnson, K. (1991). Differential effects of speaker and vowel variability on fricative perception. *Language & Speech*, *34*(3), 265-279.

Johnson, K., & Strand, E. A. (1998, January). *Speech perception without normalization or higher order invariants.* Paper presented in the "Paradigms and Problems in Phonetic Research" workshop at the annual meeting of the Lingustic Society of America, New York City.

Ladefoged, P., & Broadbent, D. E. (1957). Information conveyed by vowels. *Journal of the Acoustical Society of America*, *29*(1), 98-104.

Lass, N. J., Almerino, C. A., Jordan, L. F., & Walsh, J. M. (1980). The effect of filtered speech on speaker race and sex identifications. *Journal of Phonetics*, *8*, 101-112.

May, J. (1976). Vocal tract normalization for /s/ and /ʃ/. *Haskins Laboratories: Status Report on Speech Research*, *SR-48*, 67-73.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746-748.

Mendoza-Denton, N., & Strand, E. A. (1998). *Sociobiological ideologies of gender difference in phonetic research*. Unpublished manuscript, The Ohio State University.

Naslund, D. T. (1993). *The /s/ phoneme: A gender issue*. Unpublished manuscript, University of Minnesota, Duluth.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of vowels. *Journal of the Acoustical Society of America*, *24*, 174-184.

Remez, R. E. (1994). A guide to research on the perception of speech. In *Handbook of psycholinguistics* (pp. 145-172). San Diego, CA: Academic Press.

Schwartz, M. (1968). Identification of speaker sex from isolated, voiceless fricatives. *Journal of the Acoustical Society of America, 43*(5), 1178-1179.

Shadle, C. H. (1991). The effect of geometry on source mechanisms of fricative consonants. *Journal of Phonetics, 19*, 409-424.

Strand, E. A. (1997, October). *The role of gender stereotypes in speech perception*. Paper presented at the NWAV(E) 26 meeting, Québec City, Québec.

Strand, E. A., & K. Johnson. (1996). Gradient and visual speaker normalization in the perception of fricatives. In D. Gibbon (Ed.), *Natural language processing and speech technology: Results of the 3rd KONVENS conference, Bielefeld, October 1996* (pp. 14-26). Berlin: Mouton.

Taylor, S. E., Peplau, L. A., & Sears, D. O. (1994). *Social psychology*. Englewood Cliffs, NJ: Prentice Hall.

van Bezooijen, R. (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech, 38*(3), 253-256.

von Hippel, W., Sekaquaptewa, D., & Vargas, P. (1995). On the role of encoding processes in stereotype maintenance. *Advances in Experimental Social Psychology, 27,* 177-254.

Zebrowitz, L. A., & Montepare, J. M. (1992). Impressions of babyfaced individuals across the life span. *Developmental Psychology, 28*(6), 1143-1152.