Expressing oneself in conversational interaction

Chapter · January 2010

1 author:

Jennifer S. Pardo
Montclair State University
58 PUBLICATIONS   **1,367** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project Imitation in speech: from sensori-motor integration to the dynamics of conversational interaction View project

Project Montclair Map Task Corpus View project

11 Expressing Oneself in Conversational Interaction

Jennifer S. Pardo

A fundamental issue for speech perception is one that is common to all psychological accounts of perception, *perceptual constancy. Speech perception is complicated because a given phoneme is highly variable in its acoustic-phonetic realization, yet perception of word identity is remarkably stable*. A talker never repeats the same word in exactly the same way, and different talkers can produce the same word in markedly different ways. It is not obvious how a listener extracts a stable percept from such variable input, and understanding the process by which this is accomplished has been the primary goal of speech perception research (Liberman, 1996). Research on this problem has focused mainly on the perceptual processes involved in resolving relatively constant linguistic forms from variable acoustic realizations, with little systematic consideration of talker variability. However, acoustic-phonetic variability within and between talkers does not arise by chance—such variation is attributable to a variety of causes including talker physiology, dialect, affect, and social/situational factors. Among the important social functions that speech serves are marking a talker's identity and internal state, signaling a talker's orientation to an interactive situation, and indexing the situational roles of interacting talkers. Studies of conversational interaction demonstrate the influence of such functions on many aspects of spoken communication. A more complete understanding of this process would make the issue of linguistic variability more tractable because these sources converge on the same acoustic-phonetic dimensions, informing both speech perception and the perception of talker attributes.

Much of the past and ongoing research conducted in Robert Krauss' Human Communication Laboratory attempts to identify the patterns that nonlinguistic factors impose on the acoustic-phonetic form of utterances. In describing the current project, it is important to acknowledge the extensive contributions of Howard Giles and colleagues, whose communication accommodation theory provides a useful framework for grounding any work on acoustic-phonetic variation during conversational interaction. In addition, my favorite study by Krauss provides an important bridge to the current research (Bilous & Krauss, 1988). In our ongoing collaboration, we have found that interacting talkers engage in a subtle shift in their manner of expression, either becoming more or less similar in phonetic

repertoire. <mark>This shift provides no help for the listener in a linguistic sense, all the utterances are intelligible, but the speech forms change nonetheless.</mark>

COMMUNICATION ACCOMMODATION THEORY

The main findings of the communication accommodation theory are extensively discussed in Giles' edited book (Giles, Coupland, & Coupland, 1991). It has been found that on many occasions, a talker will converge or become more similar to an addressee in certain attributes of speech. This pattern of convergence can move up the social scale, as when a factory worker's speech becomes more similar to the speech style of a supervisor, or it can move down, as when a doctor uses speech that is tailored to a young patient. Many of the observations reported in Giles' book center on lexical choice or on sub-lexical attributes such as speaking rate. However, Giles also published two interesting empirical reports that demonstrated changes in a talker's use of accent as a function of addressee identity.

In the first study, two different interviewers engaged in conversations with the same set of Bristol talkers (Giles, 1973). One interviewer spoke using Received Pronunciation (RP), the so-called Queen's English. The other interviewer spoke with his normal Bristol accent. <mark>Excerpts of 1 min duration from the Bristol speakers' responses were played to separate Bristol listeners, who rated the degree of accentedness of the speech samples.</mark> Overall, the responses to the RP interviewer were rated as less accented than those spoken to the Bristol interviewer. These findings demonstrated that Bristol talkers converged in speaking style to the RP interviewer, becoming less Bristol-accented in their speech.

The second study introduced an interesting twist <mark>(Bourhis & Giles, 1977</mark>). A set of Welsh-born adults was asked to respond to an RP interviewer whose questions were pre-recorded. After collecting a baseline set of utterances, the Welsh speakers heard the RP interviewer make a negative statement about the fate of the Welsh language, "… the future of Welsh appears pretty dismal" (p. 125). A final set of spoken utterances was elicited with pre-recorded questions from the same RP interviewer. Two separate Welsh speakers rated excerpts from the Welsh talkers' recordings for degree of accentedness. <mark>For those Welsh talkers who exhibited strong ties to their Welsh identity, the change from the pre- to the post-insult phase of the experiment was that of divergence—these talkers' utterances became more Welsh-accented after hearing the negative comment</mark> (one talker responded by conjugating Welsh verbs). Although much of the literature has focused on convergence, it is likely that divergence has equal prevalence in many social settings.

The story of speech accommodation is complex (for an update, see <mark>Shepard, Giles, & Le Poire, 2001</mark>). Talkers converge and diverge up and down the social scale. Many aspects of a talker's identity and social situation have been found to relate to the degree to which the markers of an accent or dialect are manifested in speech (see also Labov, 1974, 1986). Although Giles' (1973) dominance measures did not correlate with his findings on accent mobility, it is reasonable to ask

whether speech variability is influenced by a very salient dimension of identity and dominance—a talker's sex.

MY FAVORITE STUDY BY KRAUSS

Krauss' Human Communication Laboratory has yielded a definitive collection of findings across a number of domains in communication. The study that has become my favorite is not cited very often relative to many of his other papers. Perhaps this is because Bilous and Krauss (1988) had an ambitious goal—to undermine the stereotype that men dominate conversational interactions with women. The findings were complex, as usual. Accommodation depended on the particular speech index that was being considered—for some indices men dominated, but for others women were dominant.

In their study, Bilous and Krauss (1988) introduced an important methodological innovation. Instead of focusing on group differences between men and women, all of the talkers first participated in a conversational task in same-sex pairs and then participated in a similar task in mixed-sex pairs. The measures of convergence/divergence were obtained by comparing the speech that was produced in same-sex pairs with the way the talkers interacted in mixed-sex pairs. The attributes they measured were total number of words, average utterance length, frequency of interruptions, frequency of short pauses (< 1 s), frequency of long pauses (> 1 s), frequency of back-channel responses, and frequency of laughter. If the male dominance hypothesis is correct, then men should not change from same-sex to mixed-sex pairings, and women should converge toward the male norm. The results showed that men and women both converged in average utterance length and frequency of short and long pauses. Women converged to male partners in total number of words and in frequency of interruptions. Men converged to female partners in frequency of back channels and frequency of laughter. However, women diverged from male partners in frequency of back channels and frequency of laughter, and men did not diverge from female partners in any of the speech indices.

This study demonstrated that accommodation is not a straightforward consequence of a talker's identity or sex—talkers converged on some speech indices at the same time that they diverged on others. This led Bilous and Krauss to conclude: "Any generalizations about the ways that men and women accommodate to each other when they interact must take into account the relevant properties of the situation in which the interaction takes place and the goals of the participants in those situations" (p. 192). Therefore, the focus of the current research is to examine the circumstances that evoke accommodation and the extent to which talkers exhibit accommodation in social settings.

A more extensive review of the literature on social interaction indicates that talkers exhibit an increase in similarity for a variety of attributes of spoken language. The findings can be organized according to the general methodology employed and the relative duration of the speech samples examined. There are a few findings that talkers who are asked to listen to spoken vowels, syllables,

or words and immediately repeat what they hear are inclined to imitate the sample items; however, the imitations are not exact in acoustic-phonetic detail (Fowler, Brown, Sabadini, & Weihing, 2003; Goldinger, 1998; Namy, Nygaard, & Sauerteig, 2002; Vallabha & Tuller, 2003; Viechnicki, 2002). These findings provide evidence that speech perception might resolve phonetic forms in enough detail to support convergence in more naturalistic social settings. Studies of social interaction that include interview situations find that talkers become more similar in attributes that are at a coarser grain than lexical tokens, including increased coordination of referential terms, description schemes, and common ground (Brennan & Clark, 1996; Clark & Wilkes-Gibbs, 1986; Garrod & Doherty, 1994; Krauss & Weinheimer, 1964). When examining longer stretches of speech, researchers have found increased similarity in sub-vocal frequency/amplitude contour, speech intensity, speaking rate, accent, and syntactic constructions (Branigan, Pickering, & Cleland, 2000; Giles et al., 1991; Gregory, 1990; Gregory & Webster, 1996; Natale, 1975). Finally, studies of dialect and accent change find that talkers become more similar to the standard of their most recent linguistic environment in attributes such as voice onset timing, vowel pronunciation, and referential term use (Labov, 1976; Sancier & Fowler, 1997). Such cumulative changes are assumed to result from individual social interactions. However, no one has provided empirical evidence that interacting talkers can become more similar to each other in sub-lexical acoustic–phonetic detail. The current project attempts to fill this void by examining changes in a talker's acoustic–phonetic repertoire before, during, and after conversational interaction.

AU: Vallabha and Tuller 2003 not in ref. list.

AU: Labov, 1976 not in ref. list.

PHONETIC CONVERGENCE

Phonetic convergence is defined as an increase in the acoustic-phonetic similarity of interacting talkers' utterances. In order to measure phonetic convergence, it is necessary to compare samples of the same lexical items spoken by each member of an interacting pair during the interaction to each other and to a set of items collected prior to the interaction. Moreover, if social settings provide the impetus for accent and dialect change, it is likely that phonetic convergence might persist beyond an individual interaction. Therefore, a corpus of speech samples was recorded from a set of talkers before, during, and after they participated in a conversational task. There were ten pairs of unacquainted talkers: six pairs of talkers formed same-sex pairs (three male and three female pairs), and four pairs of talkers formed mixed-sex pairs. To measure phonetic convergence, a combination of perceptual and acoustic measures was obtained.

In order to ensure that the talkers would produce the same lexical items during the interaction and that these items would be identifiable prior to the conversation, each pair was asked to complete six rounds of the Map Task together (developed by the Human Communication Research Center at the Universities of Glasgow and Edinburgh, Scotland; Anderson et al., 1991). The Map Task comprises paired schematic maps that contain labeled illustrated landmarks.

One member's map contains a path drawn from a starting point, around various landmarks, to a finishing point, and the companion's map contains only a starting point and landmarks. A crucial component of the Map Task is the distinction in role—one member is the instruction giver and the other is the instruction receiver. The goal of the task is for both talkers to communicate effectively enough so that the giver's path can be duplicated on the receiver's map, despite the fact that the talkers cannot see each other or each other's maps. Completion of the task requires active involvement of both participants, and spoken samples of the landmark labels can be collected both before and after the conversational interaction, to compare to those that are produced by both participants over the course of conversation.

After all pairs of talkers had completed five or six pairs of maps together, the recordings were coded for between-talker repetitions of the same landmark label phrases. These task repetitions and a set of pre- and post-task versions of the same lexical items were excised from the recordings to create a sensitive test of phonetic convergence. The perceptual assessment of phonetic convergence comprised a listening test that used an AXB design. On each trial, a listener heard three samples of the same lexical items. The middle or X-item was always an utterance of the phrase produced by one of the two members of a pair. The A and B items were always repetitions of the phrase produced by the other member of a pair. The task for a listener was to choose which item, A or B, sounded more like the middle item in terms of its pronunciation. On some of the trials, the flanking items to be compared to the middle item included the repetition from the task session and the pre-task session item. On other trials, the flanking items comprised the pre-task session item and the post-task session item. The order of presentation of the items was counterbalanced so that all items had a 50% chance of being presented in A or B.

If a listener judged the task repetition to be more similar in pronunciation to the partner's task sample item than the pre-task session item, then the talker has converged in pronunciation from the pre-task session to the task session. If a listener judged the post-task session version to be more similar in pronunciation than the pre-task session item, then the talker has persisted in phonetic convergence beyond the task session. If a listener is unable to express a preference, then the responses will occur randomly across the trials, leading to a 50% preference for the task repetition or the post-task session item. Therefore, the perceptual assessment of phonetic convergence can be quantified as a greater than chance (50%) judgment of the task repetition or post-task session items as more similar in pronunciation than the pre-task session items. Moreover, the degree of phonetic convergence was compared across the factors of *Timing* in the Map Task session (early vs. late), *Talker Role* (giver vs. receiver), and *Pair Sex* (male vs. female, or male givers vs. female givers). Thirty listeners provided judgments for the same-sex pairs, and twenty-one listeners judged samples from the mixed-sex pairs of talkers.

Overall, the AXB perceptual assessments of phonetic convergence confirmed that paired talkers became more similar to each other in their pronunciation of

repeated landmark label phrases during performance of the Map Task. For the same-sex pairs, listeners judged the task repetition or post-task session items to be more similar on 62% of the trials. Mixed-sex pairs also yielded a significant effect, with listeners finding the key items more similar on 53% of the trials (both measures were statistically different from chance, 50%, p < 0.05). However, the AXB protocol for the mixed-sex pairs required that a listener choose among items produced by a female talker in comparison to an item from a male talker and vice versa. Perceptual comparisons between male and female speech are particularly difficult, and may have affected the ability of listeners to detect phonetic convergence with mixed-sex pairs of talkers. Alternatively, it could be the case that men and women do not converge as readily as they do when in same-sex pairs, but this speculation awaits further research.

Phonetic convergence is not an all-or-none phenomenon. A number of factors influence it. For the same-sex pairs, convergence was found to occur relatively early (62%), increase in the second half of the conversations (68%), and persist into the post-task session (62%). Overall, receivers converged more than givers. When givers provided the sample items repeated at short delay by receivers, the receivers' repetitions showed phonetic convergence (62%), but not as great as when the receivers provided the sample items that a giver repeated (68%). Furthermore, female pairs of talkers converged less than male pairs of talkers (58% female; 75% male). Finally, the role of the talker interacted with the sex of the pairs—female receivers did not converge to female givers (47%), female givers converged to female receivers (62%), male receivers converged to male givers (71%), and male givers converged to male receivers (64%). These data and analyses are reported in much more detail in Pardo (2006). For mixed-sex pairs, there were similar trends in the data, but the effects were not as pronounced, and some cells failed to reach significance. The effect of role showed an interesting pattern. Male receivers also failed to converge to female givers (51% n.s.), but female receivers converged to male givers (55%). Taken together, there is no evidence that receivers converged to female givers, whether in same-sex or mixed-sex pairings. These findings echo those found by Bilous and Krauss (1998)—accommodation is influenced by a talker's sex in ways that are not attributable to hypothetical male dominance. Furthermore, these findings provide the first indication that phonetic convergence is also influenced by the role that a talker has in a structured task situation.

ACOUSTIC MEASURES OF CONVERGENCE

Although a listener under these circumstances makes a global judgment of an utterance based on multiple acoustic-phonetic dimensions, it is useful to ask whether any single acoustic dimension is driving these judgments. For example, pitch and speaking rate are relatively salient attributes of speech. If listeners were basing their judgments of similarity on fundamental frequency (F0, roughly heard as pitch) or utterance duration (related to speaking rate), then the variability in the AXB judgments should be related to the variability in these acoustic attributes. To determine the extent that variability in the AXB data are accounted for by variability in the

degree to which a task repetition or post-task session item matched a task sample item in average F0 and/or duration, multiple regression analyses treated these factors as predictors of the AXB data from the same-sex talker pairs.

First, the acoustic measures were converted into measures reflecting the acoustic similarity between a task sample item and its task repetition or post-task session item—the absolute values of the differences for each pair of label phrases in average F0 and duration were computed for each AXB comparison type. These data were entered into the multiple regression analysis, which found that the linear combination of average F0 and duration was related to the AXB similarity test data, accounting for 28% of the variance in the AXB scores (adjusted $R^2 = .279$, $F(2,69) = 14.756$, $p < .001$). Moreover, each acoustic factor predicted the AXB similarity data over and above the other factor, with average F0 accounting for 12% and duration 20% of the variance in the AXB similarity data (average F0: adjusted $R^2 = .117$, $F(1,70) = 10.363$, $p < .002$, R^2 change $= .085$ $F(1,69) = 8.350$, $p < .005$; duration: adjusted $R^2 = .204$, $F(1,70) = 19.151$, $p < .001$, R^2 change $= .171$, F change $(1,69) = 16.809$, $p < .001$; correlation between duration and average F0, $r = .154$, $p = .098$). Although average F0 and duration were linearly related to listeners' judgments of similarity, there is a great deal of variability in the AXB data that is not accounted for by these acoustic measures.

Because there was an interaction between the effects of Talker Role and Pair Sex in the AXB similarity data, analogous multiple regression analyses were performed separately for female and male talkers. For female talkers, the linear combination of average F0 and duration was related to the AXB data, accounting for 41% of the variance, but not for male talkers, for whom the acoustic measures were only marginally related to the AXB scores [females: adjusted $R^2 = .411$, $F(2,33) = 13.210$, $p < .001$; males: adjusted $R^2 = .069$, $F(2,33) = 2.295$, $p = .117$]. In separate regression analyses for each acoustic factor, the only significant finding was that duration predicted the AXB scores for the female talkers over and above average F0 [duration: adjusted $R^2 = .413$, $F(1,34) = 25.611$, $p < .001$, R^2 change $= .351$, $F(1,33) = 20.851$, $p < .001$; average F0: adjusted $R^2 = .067$, $F(1,34) = 3.516$, $p = .069$, R^2 change $= .015$ $F(1,33) = .891$, $p = .4$ n.s.; males all adjusted $R^2 < .03$, n.s.].

Overall, the regression analyses found that perceived phonetic convergence is somewhat related to similarity in duration and is weakly related or unrelated to similarity in average F0. However, the relationship appears to be obtained only for female talkers. Because similarity in duration did not predict variability in male AXB data (nor perfectly for the female AXB data), there must be additional dimensions that listeners use when judging similarity, and possibly there are different standards applied to male and female talkers.

ANALYSES OF PRE-TASK AND POST-TASK VOWEL SPACES

The acoustic analyses of the items from the conversational task indicated that unidimensional acoustic parameters, such as average F0 or item duration, bore a small to moderate relationship to perceivers' judgments for female talkers. The

next set of acoustic measures assessed whether the persistence of phonetic convergence could be due to changes in vowel pronunciation and estimated the extent to which changes in a talker's vowel space aligned with an individual partner and a talker's role. Because vowels are exposed to a variety of influences, a talker's vowel repertoire is a likely domain of phonetic convergence. These influences include changes in resonant frequencies related to speaking register, dialect, and ideolect.

In order to characterize acoustic variability among vowels, early research on the production and perception of speech (e.g., Peterson & Barney, 1952) settled on a two-dimensional representation using the lowest resonant components, the first and second formants (F1 and F2). This representation also follows linguists' classifications of the vowels as high versus low (jaw height, roughly the inverse of F1) and back versus front (tongue position, roughly F2). The vowels in *heat* and *hoot* (\i\ and \u\) are high (with low F1), and the vowels in *hat* and *hot* (\œ\ and \A\) are low (with high F1); \i\ and \œ\ are more fronted than \A\ and \u\ (having higher vs. lower F2). In general, these four point vowels sit at the periphery of the vowel space, with other vowels in intermediary steps between them. Vowels are highly variable acoustically, and perceptually, vowels yield a greater degree of within-category discrimination than consonants (Pisoni, 1973, 1975). On these grounds, vowels are likely candidates for convergence—there is both acoustic room for and perceptual resolution of variability among vowel tokens. Moreover, Labov (1974, 1986) has described vowels as central to sound change and dialectal differences among talkers of American English.

Recent evidence suggests that a talker cannot match his or her own vowel productions exactly, with mimicked productions falling to the center of sample items (Viechnicki, 2002) or following idiosyncratic biases in variation (Vallabha & Tuller, 2004). Therefore, exact matching of vowel spectra is not expected either within or between talkers, but partners may still converge in vowel formant frequencies. The vowel measures were taken from subsets of the speech collected in the pre-task and post-task sessions. For each session, three repetitions of the point vowels, \i\, \œ\, \A\, and \u\, each embedded in hVt words in the carrier sentence, "Say again," comprised materials for acoustic analysis. The vowel sets were digitally analyzed using Praat, a freeware acoustic analysis software package designed by Paul Boersma and David Weenink at the Institute of Phonetic Sciences at the University of Amsterdam. Measures of F1 and F2 were taken at the midpoint of each vowel token from linear prediction spectrographic analyses. Figure 11.1 illustrates the F1 × F2 vowel spaces in the pre-task and post-task sessions for all of the talkers in the same-sex pairs. The acoustic analyses of vowel space indicated that none of the talkers matched average F1 × F2 vowel spectra from the pre-task to the post-task session. Statistical analyses of vowel spectra found that there were differences in F1 and F2 from the pre-task to the post-task sessions for both male and female talkers, and indicated that these differences were influenced by a talker's role and vowel identity.

To quantify the differences in vowel dispersion more precisely with respect to phonetic convergence, the acoustic measures were converted to Euclidean
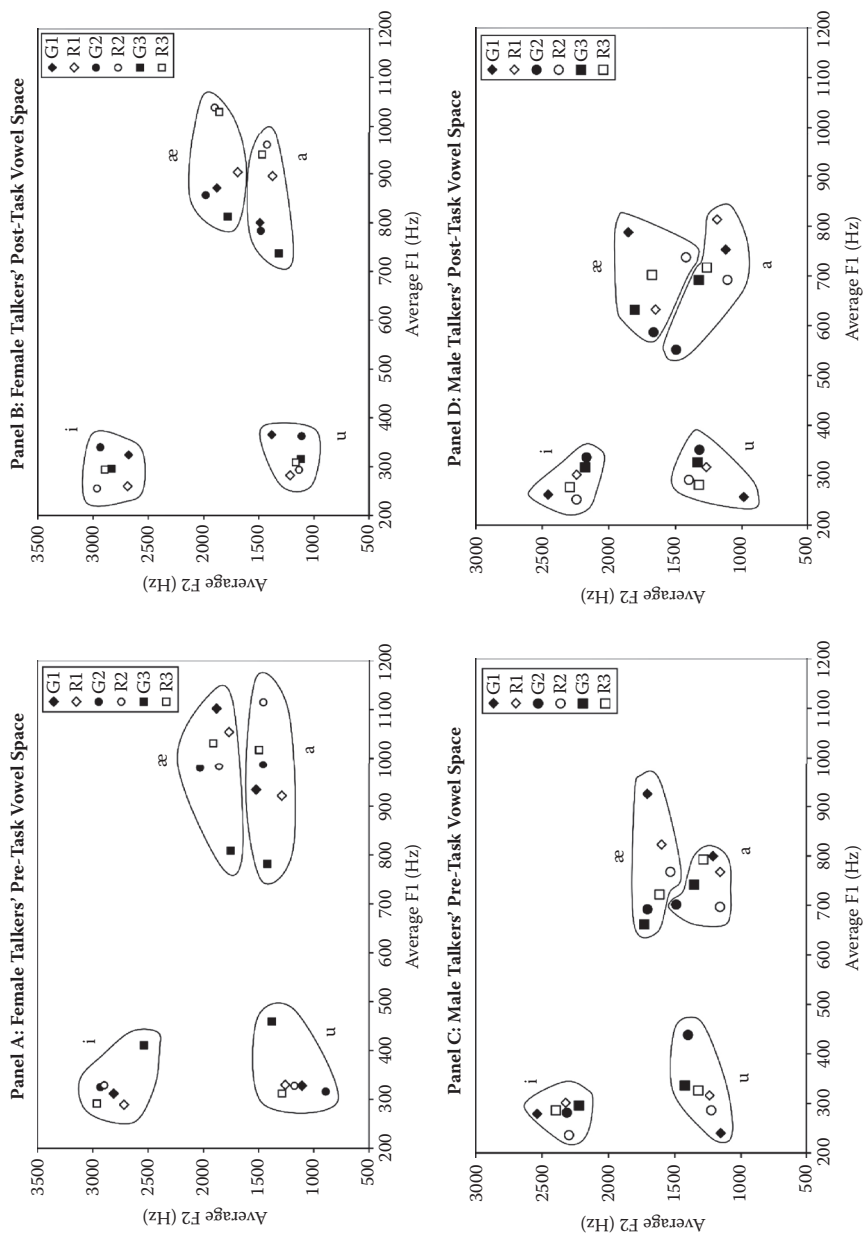
FIGURE 11.1(A–D) Plots of the F1 × F2 vowel spaces for same-sex pairs of talkers. Panels A and B comprise pre-task and post-task vowel measures for female talkers, and panels C and D contain the same measures for male talkers. Filled bullets correspond to givers and open bullets to receivers. Bullets of the same shape correspond to talkers who were paired in the intervening Map Task session.
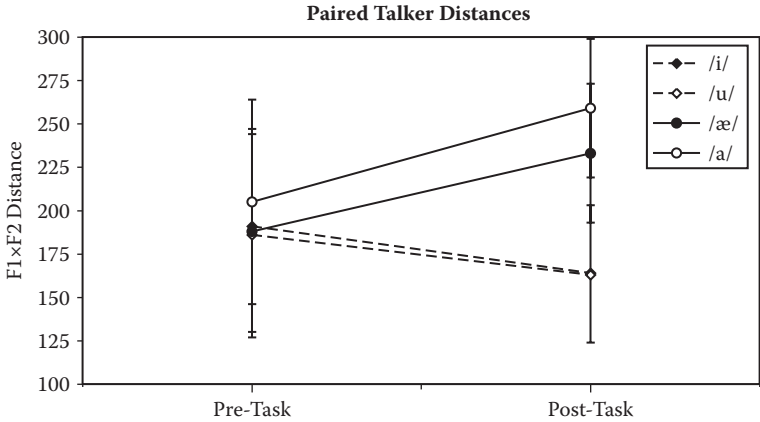
FIGURE 11.2 Vowel convergence measures for paired talkers with 95% confidence intervals. Each point corresponds to the average of the inter-talker distances for the same vowel in the pre-task and post-task sessions.

distances in the F1 × F2 space. The first set of distance measures assessed phonetic convergence by comparing pre-task differences between paired talkers with post-task differences for all four vowels. Figure 11.2 displays the inter-talker distance measures for the pre-task and post-task sessions across the point vowels, showing that paired talkers converged in \i\ and \u\ and diverged in \œ\ and \ʌ\. These inter-talker distance measures were analyzed in a mixed design analysis of variance (ANOVA) for the within items effects of *Session* (pre-task vs. post-task), *Sex* (female vs. male), and *Pair* (1–3), and the between items effect of *Vowel* (\i\, \œ\, \ʌ\, and \u\). The mixed design ANOVA confirmed the interaction between *Session* and *Vowel* ($F(3,8) = 5.346$, $p < .027$), and these factors also interacted with *Pair* and *Sex*, verifying the earlier observations of unique convergence patterns for different pairs of talkers ($F(6,16) = 4.247$, $p < .011$). The inter-talker distance measures demonstrated that talkers both converged and diverged in distinct aspects of their phonetic repertoire.

The measures of vowel centralization assessed the tendency for each talker to produce more central vowels by calculating distances between the front vowels (\i\ to \œ\) and the back vowels (\u\ to \ʌ\) in the pre-task and post-task sessions. The talkers centralized vowel spectra from the pre-task (d = 786) to the post-task (d = 728) session, and givers' vowel spectra (d = 694) were more central than receivers' (d = 820). More importantly, as shown in Figure 11.3, givers centralized vowel spectra from the pre-task to the post-task sessions more so than receivers did. A mixed design ANOVA on the centralization measures for the within items effects of *Session* (pre-task vs. post-task), *Talker Role* (giver vs. receiver), *Sex* (female vs. male), and *Pair* (1–3), and the between items effect of *Comparison Type* (\i\ to \œ\ and \u\ to \ʌ\) confirmed the main effect of *Session* ($F(1,4) = 33.054$, $p < .006$) and the main effect of *Talker Role* ($F(1,4) = 119.470$, $p < .001$) and
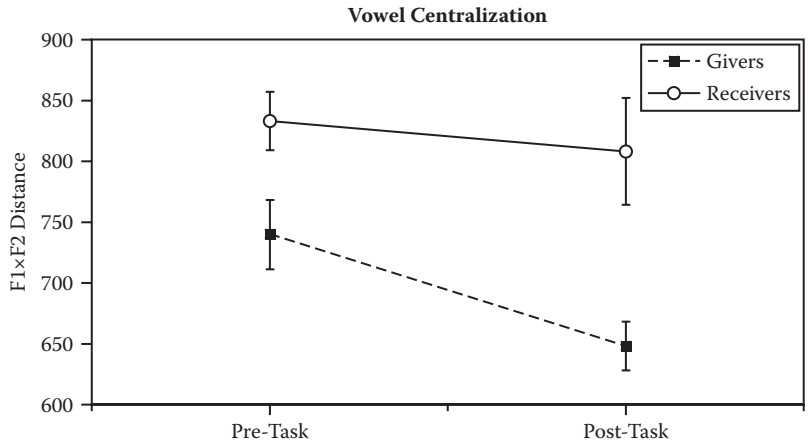
FIGURE 11.3 Vowel centralization measures for high versus low vowels with 95% confidence intervals. Each point corresponds to the average of the inter-vowel distances (\i\ to \œ\ and \u\ to \ʌ\) for the givers (filled bullets) and receivers (open bullets) in the pre-task and post-task sessions.

the interaction between *Session* and *Talker Role* ($F(1,4) = 10.567$, $p < .04$); the three-way interaction with *Sex* was not significant (p = .9 n.s.). Despite the fact that all talkers centralized from the pre-task to the post-task sessions and givers' vowels were more centralized overall than those of receivers, givers exhibited greater centralization across the pre-task and post-task sessions than receivers. Combined with the finding that the high vowels converged between talkers and the low vowels diverged, it appears that givers differed from receivers mainly in the centralization of the low vowels, \œ\ and \ʌ\.

Taken together, it appears that givers diverged from receivers in the low vowels, \œ\ and \ʌ\, and these vowels centralized more for givers than for receivers. Although definitive proof is difficult to establish, the measures of vowel spectra seem to indicate a shift toward centralization for givers' low vowels. Receivers, on the other hand, largely maintained the peripheral expression of their low vowels. This phenomenon could be a starting point for the kinds of shifts in vowels that Labov (1974, 1986) described for dialectal changes, in which the back vowels shift to more fronted positions in the vowel space. The high vowels showed some convergence, but again, this has not been demonstrated unequivocally. These measures indicate that talkers differ in their use of vowel variants, and the pattern may be consistent with their role in conversational interaction. These findings of co-occurring convergence and divergence indicate that some attributes are free to vary with the partner, but others are reserved for individual expression, perhaps to mark a talker's approach to his or her role. Because vowel centralization is often related to a casual speech setting, it is possible that givers were attempting to soften their distinct role by signaling a casual attitude at the same time that they converged to their partners in other regions of the vowel space.

CONCLUSION

This project establishes the phenomenon of acoustic-phonetic convergence in a spontaneous social setting. Convergence occurred early, increased over the course of an interaction, and persisted beyond the conversational setting. More importantly, the social constraints of role and talker sex modulated the degree of acoustic-phonetic convergence. Acoustic measures demonstrated that these subtle changes are not universal—the perceptually salient acoustic dimensions of pitch and speaking rate were mostly unrelated to perceivers' global assessments of phonetic convergence, and some vowels converged at the same time that other vowels diverged. A more complete understanding of the extent, means, and purpose of acoustic-phonetic convergence awaits future investigations. This phenomenon does not appear to serve any linguistic aims; indeed, all utterances were intelligible and the talkers performed the task with relative ease. It is likely that these changes served other nonlinguistic purposes, as proposed so eloquently by Krauss and Chiu (1998):

> Linguists regard language as an abstract structure that exists independently of specific instances of usage…, but any communicative exchange is situated in a social context that constrains the linguistic forms participants use. How these participants define the social situation, their perceptions of what others know, think, and believe, and the claims they make about their own and others' identities will affect the form and content of their acts of speaking. (p. 41)

ACKNOWLEDGMENTS

REFERENCES

Anderson, A. H., Bader, M., Bard, E. G., Boyle, E., Doherty, G., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., Sotillo, C., Thompson, H. S., & Weinert, R. (1991). The HCRC Map Task corpus. *Language & Speech, 34,* 351–366.

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

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.

Bradlow, A. R. (2002). Confluent talker- and listener-oriented forces in clear speech production. In C. Gussenhoven & N. Warner (Eds.), *Laboratory phonology 7* (pp. 241–273). New York: Mouton de Gruyter.

Au: Bradlow 2002 not cited in text.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition, 75*, B13–25.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 22*, 1482–1493.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition, 22*, 1–39.

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

Garrod, S., & Doherty, G. (1994). Conversation, co-ordination and convention: An empirical investigation of how groups establish linguistic conventions. *Cognition, 53*, 181–215.

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

Giles, H., Coupland, J., & Coupland, N. (1991). *Contexts of accommodation: Developments in applied sociolinguistics*. Cambridge University Press.

AU: Need location of publisher.

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

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

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

Krauss, R. M., & Chiu, C-Y. (1998). Language and social behavior. In D. Gilber, S. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* Vol. 2 (4th ed., pp. 41–88) Boston: McGraw-Hill.

Krauss, R. M., & Weinheimer, S. (1964). Changes in the length of reference phrases as a function of social interaction: A preliminary study. *Psychonomic Science, 1*, 113–114.

Labov, W. (1974). Linguistic change as a form of communication. In A. Silverstein (Ed.), *Human communication: Theoretical explorations* (pp. 221–256). Hillsdale, NJ: Lawrence Erlbaum Associates.

Labov, W. (1986). Sources of inherent variation in the speech process. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in the speech processes* (pp. 402–425). Hillsdale, NJ: Lawrence Erlbaum Associates.

Liberman, A. M. (1996). *Speech: A special code*. Cambridge, MA: The MIT Press.

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

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

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

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral & Brain Sciences, 27*, 169–190.

AU: Pickering and Garrod 2004 not cited in text.

Sancier, M. L., & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics, 25*, 421–436.

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.

Vallabha, G. K., & Tuller, B. (2004). Perceptuomotor bias in the imitation of steady-state vowels. *Journal of the Acoustical Society of America, 116*, 1184–1197.

Viechnicki, P. D. (2002). Composition and granularity of vowel production targets. *Dissertation Abstracts International. 63*(4-A), 1320.