

# Phonetic convergence in college roommates

Alexandra Suppes, Robert Krauss


*Journal of Phonetics*

## Cite this paper

Downloaded from [Academia.edu](#) 

[Get the citation in MLA, APA, or Chicago styles](#)

## Related papers

[Download a PDF Pack](#) of the best related papers 



[Conversational role influences speech imitation](#)

Isabel Jay

[Phonetic imitation of L2 vowels in a rapid shadowing task](#)

Arkadiusz Rojczyk

[Phonetic convergence in spontaneous conversations as a function of interlocutor language distance](#)

Midam Kim



## Phonetic convergence in college roommates

Jennifer S. Pardo<sup>a,\*</sup>, Rachel Gibbons<sup>b</sup>, Alexandra Suppes<sup>c</sup>, Robert M. Krauss<sup>b</sup>

<sup>a</sup> Department of Psychology, Montclair State University, 1 Normal Avenue, Montclair, NJ 07043, United States

<sup>b</sup> Department of Psychology, Columbia College, Columbia University, United States

<sup>c</sup> Department of Public Health, Weill Cornell Medical College, United States

### ARTICLE INFO

#### Article history:

Received 20 March 2010

Received in revised form

22 September 2011

Accepted 1 October 2011

Available online 19 October 2011

### ABSTRACT

Previous studies have found that talkers converge or diverge in phonetic form during a single conversational session or as a result of long-term exposure to a particular linguistic environment. In the current study, five pairs of previously unacquainted male roommates were recorded at four time intervals during the academic year. Phonetic convergence over time was assessed using a perceptual similarity test and measures of vowel spectra. There were distinct patterns of phonetic convergence during the academic year across roommate pairs, and perceptual detection of convergence varied for different linguistic items. In addition, phonetic convergence correlated moderately with roommates' self-reported closeness. These findings suggest that phonetic convergence in college roommates is variable and moderately related to the strength of a relationship.

© 2011 Elsevier Ltd. All rights reserved.

### 1. Introduction

The acoustic–phonetic form of a word varies widely both between and within talkers. Production of the same word across talkers differs according to anatomy, sex, age, dialect, and region of residence. In contrast, variability in an individual talker's production of a word on different occasions is less noticeable in everyday conversation. Much of this variability can be attributed to semantic and pragmatic impact on usage. However, talkers have also been found to vary acoustic–phonetic form with very recent exposure to another talker and after prolonged exposure to a particular linguistic environment. In particular, talkers have been found to become more similar in acoustic–phonetic form to a model or to an ambient linguistic environment, exhibiting phonetic convergence or gestural drift (e.g., Babel, 2010; Evans & Iverson, 2007; Goldinger, 1998; Namy, Nygaard, & Sauerteig, 2002; Pardo, 2006; Sancier & Fowler, 1997). Missing from the literature is an understanding of the dynamics of phonetic convergence in a pair of talkers who interact for longer than a single experimental session. Thus, the current study examined phonetic convergence in previously unacquainted college roommates through the academic year and related measures of phonetic convergence to measures of perceived closeness.

#### 1.1. Speech accommodation

Communication Accommodation Theory proposes that individuals use language to achieve a desired social distance between the

self and interacting partners (Shepard, Giles, & Le Poire, 2001). Accordingly, convergence refers to the ways in which a talker adjusts speaking style to become more similar to an interacting partner, whereas divergence refers to changes in speaking style that result in reduced similarity to a partner. The changes initially observed in speech included attributes measured over long stretches of dialog such as accent, speaking rate, intensity, low-frequency band variation, pause frequency, and utterance length (e.g., Giles, Coupland, & Coupland, 1991; Gregory, 1990; Gregory, Dagan, & Webster, 1997; Gregory & Webster, 1996; Natale, 1975). The reasons proposed for employing accommodation are varied, but the most prevalent is the similarity attraction hypothesis, which claims that individuals try to be more similar to those to whom they are attracted (Byrne, 1971). This proposal has evoked many hypothetical functions for convergence–convergence could result from a need to gain approval from the interacting partner (Street & Giles, 1982), from a concern that the interaction is carried out smoothly (Gallois, Giles, Jones, Cargile, & Ota, 1995), or from an effort to increase one's own intelligibility during the interaction (Triandis & Triandis, 1960). Divergence, on the other hand, can be used to accentuate individual differences or to display disdain for another individual (Bourhis & Giles, 1977; Shepard et al., 2001).

Although Communication Accommodation Theory was developed in the context of studies that employed social settings, convergence and mimicry have also been examined in more restricted laboratory settings. For example, talkers who were asked to repeat recorded words sampled from another talker produced utterances that were more similar to those of the sample talker than their baseline utterances (Goldinger, 1998; Namy et al., 2002; Shockley, Sabadini, and Fowler, 2004; but see Vallabha & Tuller, 2004). In these studies, talkers were first recorded producing

\* Corresponding author. Tel.: +1 973 655 7924.

E-mail addresses: [pardoj@mail.montclair.edu](mailto:pardoj@mail.montclair.edu), [pardo@optonline.net](mailto:pardo@optonline.net) (J.S. Pardo).

a baseline series of words prompted by a list, and then were instructed to listen to the same series of words produced by another talker and to repeat each word immediately. Imitation of word forms was assessed by asking a separate set of listeners to judge the similarity of the baseline and the shadowed utterances to the model utterances. Across multiple studies using this technique, shadowing talkers were found to converge (i.e., imitate or become more similar to) the model talkers that they heard. [Shockley et al. \(2004\)](#) also related perceived convergence to variation in an acoustic–phonetic attribute (voice onset time: VOT). Thus, shadowers were not only heard as sounding more similar to the model, but they also increased VOTs of shadowed tokens when shadowing a model token that had a lengthened VOT.

In a study that examined phonetic variation with long-term exposure to different linguistic environments, [Sancier and Fowler \(1997\)](#) recorded a native Brazilian Portuguese and English (L2) speaker on three occasions: once after spending four months in the US before leaving for Brazil, again just after returning to the US after spending 2.5 months in Brazil, and once more after spending four months in the US. They found that the talker's Brazilian Portuguese utterances were judged by Brazilian Portuguese listeners to be more accented after the talker had spent four months in the US compared to utterances produced after spending 2.5 months in Brazil. In addition, the talker's VOTs in both Brazilian Portuguese and English had lengthened after her stays in the US, converging toward the average VOTs in English. This finding was surprising because adjustments were made in the language that the talker had not been using while in the US, indicating a process that is super-ordinate to a particular language.

### 1.2. Vowel convergence

Other studies have extended phonetic convergence to measures of vowel formants ([Babel, 2009, 2010](#); [Evans & Iverson, 2007](#); [Pardo, 2010](#); [Pardo, Cajori Jay, & Krauss, 2010](#)). For example, [Babel \(2009, 2010\)](#) has assessed variation in the first and second formants of vowels in shadowing tasks. Overall, inter-talker distances in vowel formants were reduced from baseline to shadowed tokens for some of the vowels, and the degree of adjustment was related to the talkers' implicit attitudes toward the race or nationality of the model talker. In a longer-term study of accent change similar to that of [Sancier and Fowler \(1997\)](#), [Evans and Iverson \(2007\)](#) reported that Northern British English students shifted pronunciation of some of their vowels as a result of spending up to two years at school in Southern England. After only three months, some of the talkers had started to shift some of their vowels, but after one and two years, most of the talkers had adopted the shifted dialectal variants. Moreover, their rated degree of southern accentedness also became stronger over the course of two years in school.

In both approaches, vowels were found to change, but the only indication that the changes were perceptible derives from ratings of accentedness, which relates to global dialect convergence rather than individual phonetic convergence. In two studies of phonetic convergence during conversational interaction, [Pardo \(2010; Pardo et al., 2010\)](#) found that talkers converged on some vowels while diverging or not changing on other vowels, and that the degree of vowel convergence/divergence was related to the role of the talker. Furthermore, Pardo et al. found that the degree of vowel convergence (measured as reduction in inter-talker distances) was moderately related to the perceived convergence of receivers to givers ( $r(10) = -0.59$ ,  $p = 0.04$ ). However, there were other patterns of perceived convergence that were not readily attributable to vowel formants or to articulation rates, which indicates that perceived phonetic convergence is likely to result from multidimensional impressions of phonetic variation.

### 1.3. Conversational convergence

The ability to converge to another talker's word pronunciation or to a linguistic environment suggests detailed perceptual resolution and closely coupled perception and production. However, it is necessary to delineate the factors that modulate this process in more natural settings of language use, such as during a conversational interaction ([Pardo, 2006, 2010](#); [Pardo et al., 2010](#)). Although interacting talkers have been found to converge in acoustic–phonetic form, the degree of convergence was subtle and was consistently influenced by the sex of the pair of talkers and a talker's role in the conversation. Indeed, talkers converged on some acoustic–phonetic dimensions at the same time that they diverged on others ([Pardo, 2010](#); [Pardo et al., 2010](#); see also [Bilous & Krauss, 1988](#)). Therefore, phonetic convergence is not an automatic consequence of detailed perceptual resolution, but has variable effects on different speech attributes and is unconsciously modulated by each talker's interpretation of the situation.

At present, accommodation has been established as a prevalent yet variable phenomenon in the speech of interacting talkers ([Giles et al., 1991](#); [Shepard et al., 2001](#)). With respect to dialect acquisition and change, [Labov \(1986\)](#) has found that individuals employ different vowel variants across different social settings, and that the use of local dialect markers is related to an individual's attitudes toward the area ([Labov, 1972](#); see also [Babel, 2010](#); [Eckert, 1989](#)). Moreover, because the location of dialect boundaries coincides with geographical boundaries that reduce opportunities for direct social interaction, [Labov \(1974\)](#) has proposed that dialect variation and change result from opportunities for direct social interaction. Yet, despite an explicit desire to accommodate to an ambient linguistic environment, long-term opportunities for social interaction with native speakers, and an unconscious tendency to imitate speech, most talkers fail to eradicate a foreign accent or to lose all markers of a regional dialect. Even though the talkers in previous studies of long-term change made adjustments that were perceptible as changes in relative accentedness, they never sounded fully Southern or unaccented ([Evans & Iverson, 2007](#); [Sancier & Fowler, 1997](#)). In order to understand the limitations of these processes, it is necessary to examine phonetic convergence at the level of individual pairs of talkers. To date, the dynamics of linguistic variation that result from continued contact with the same individual have not been studied empirically.

### 1.4. The current study

The present study attempts to fill this gap by examining phonetic convergence among talkers who interact on a daily basis, college roommates. In order to assess phonetic convergence, speech samples were collected from previously unacquainted college roommates at four intervals during the academic year and were used to elicit measures of perceptual similarity from a separate set of listeners. In addition, measures of item duration and vowel spectra were collected and compared to perceptual assessments of phonetic convergence. According to Communication Accommodation Theory and findings of phonetic convergence during conversational interaction, the roommates ought to exhibit phonetic convergence relatively early in the academic year. If phonetic convergence follows a similar trajectory to the gestural drift observed by [Sancier and Fowler \(1997\)](#), then convergence should increase prior to winter break and decrease after the roommates return from winter break. However, the current study is being conducted on individuals who do not undergo cross-language alternation, so the roommates might not show a decrease in convergence after returning from winter break. If vowel

convergence follows a trajectory similar to that reported by [Evans and Iverson \(2007\)](#), then only some of the talkers should demonstrate vowel convergence after the first three months of cohabitation. However, because the current study examined vowel convergence in roommates as opposed to dialectal changes in vowel formants, it is possible that vowel convergence might emerge earlier than expected in the current study. In addition, if phonetic convergence is related to a talker's attitude toward their roommate, then the degree of phonetic convergence should be related to the roommates' reported feelings of closeness, which was measured at the end of the first semester ([Babel, 2009, 2010; Gallois et al., 1995; Street & Giles, 1982](#)).

## 2. Method

### 2.1. Participants

A total of 10 male Columbia College undergraduates (5 pairs of roommates, aged 19–21) provided speech samples. All talkers were native English speakers with no reported hearing or speech disorders. The talkers received compensation at a rate of \$10 per hour. [Table 1](#) displays the place of origin for each talker in each pair.

A total of 30 members of the Columbia University undergraduate population provided perceptual similarity judgments of excerpts from the recordings. All listeners were native English speakers with no reported hearing or speech disorders. The listeners received course credit in exchange for completion of the task.

### 2.2. Materials

#### 2.2.1. Recordings

At all time intervals, the roommates provided 5 sets of American English vowels embedded in hVd/t words in the carrier sentence, *Say \_\_\_ again* (mixed with filler items). In addition, each talker provided two utterances of two sentences prompted by printed sheets, *She had your dark suit in greasy wash water all year* and *Don't ask me to carry an oily rag like that*. These sentences were chosen because they include phrases that are phonologically diverse, that include the four point vowels, and that exhibit variation across US dialect regions (*dark suit, greasy, oily rag, and wash water*). For example, [Clopper and Pisoni \(2004\)](#) reported that New England talkers were more likely than others to produce an *r*-less form of *dark* and Southern talkers were more likely to produce a voiced fricative in *greasy*.

#### 2.2.2. Relationship quality

In order to assess the quality of the roommates' relationship, they completed a Roommate Relationship questionnaire at the end of their first semester together. The questionnaire included requests to estimate the amount of time spent with their roommate throughout the first semester, the number of hours per week the two spent in the same room, and the number of meals eaten together per week. Other questions asked about the quality of their relationship and how much they liked their roommate. Specific measures included the closeness they felt toward their roommate (on a 7-point scale), the amount in common they felt they shared with their roommate (on a 7-point scale), and the Inclusion of Other in the Self Scale (IOS; [Aron, Aron, & Smollan, 1992](#)). The IOS uses a series of seven paired circles that range from showing no overlap to showing almost complete overlap. Each talker selected the circle combination depicting the overlap that best represents the closeness they feel to their roommate.

#### 2.2.3. Similarity test

To compose items for the similarity test, four key phrases were excised from the sentences produced at each time period: *dark suit, greasy, oily rag, and wash water*. The second repetition of each item was chosen in all but two cases, in which the first utterance was used (a cough interrupted the second repetition of one word phrase, and a mispronunciation affected another).

### 2.3. Procedure

#### 2.3.1. Recordings

Roommate pairs were recruited during the summer before they would spend an academic year living together. Using a list provided by Columbia University Housing Services, the study recruited only those roommate pairs who did not enter the housing lottery together, typically indicating that they had no prior relationship. At the time of their first recording, the talkers were further questioned about any prior contact with their roommates, and none of the pairs had communicated previously.

Speech samples were collected at four time intervals throughout the academic year during which the roommates cohabitated: T1 recordings were sampled in late August, before the roommates had met; T2 recordings were sampled in late October; T3 recordings were sampled in December at the end of the first semester and just prior to winter break; and T4 recordings were sampled in January when the roommates returned from winter break, before they resumed interaction. Both members of each pair completed a recording session individually at each time point. The questionnaire assessing the quality of the roommates' relationship was administered after the recording session at T3 in December. One pair did not return for the T4 recording session, therefore, their phonetic convergence measures from T2 and T3 were excluded from the statistical analyses, but were used to relate phonetic convergence at T3 to closeness ratings taken at T3. The recordings were obtained via a head-mounted AKG microphone connected to a Superscope PSD300 CD digital recorder.

#### 2.3.2. Perceptual tests

To assess phonetic convergence, an AXB similarity test was presented to a separate set of listeners. The listening test comprised trials that were designed to assess phonetic convergence at each time interval by asking a listener to judge similarity in pronunciation between the roommates' speech at T2 (October), T3 (December), and T4 (January), relative to their pronunciation at T1 (August). On each trial, a listener heard three repetitions of the same word or phrase in which an utterance from one talker (X) that was produced at either T2, T3, or T4 was flanked by two versions of the same phrase spoken by their roommate (A and B). One of the flanking items was the baseline utterance produced at T1, before the talkers had interacted, and the other flanking item was the corresponding utterance from the relevant time interval (T2, T3, or T4). On half the trials, utterances from one member of a roommate pair were used as X-items, and the other half of the trials used the utterances from the other member of a roommate pair as X-items. The order of presentation of T1 and T2/T3/T4 items was counterbalanced so that T1 items were presented in position A on half of the trials and in position B on the other half of the trials. Thus, at the second time interval, if the roommates sounded more similar to each other than they had prior to meeting, the T2 item of one talker should be chosen as sounding more like the T2 item of his roommate than that talker's T1 item.

In order to assess phonetic convergence, listeners were asked to determine which of the two flanking items (A or B) sounded more similar to the middle item (X) in terms of its pronunciation. Each trial began 1000 ms after a listener indicated a response,

and the items in each trial were presented at 200 ms ISI. The presentation of trials was blocked by roommate to keep the speaker in the X position the same throughout a block. The AXB test was presented over Sennheiser HD 280 Pro headphones connected to Macintosh computers running PsyScope, and responses were collected via keyboard using the 1 (first item) and 0 (last item) keys.

### 3. Results

#### 3.1. Perceptual assessment

If phonetic convergence between roommate pairs occurred, then the utterances produced by both roommates at later time intervals should sound more similar to each other than the baseline utterances produced at T1. Therefore, responses to the AXB test trials were scored as the percentage of trials on which a later time interval utterance (T2, T3, or T4) was judged to be more similar to the X-item (the roommate's T2, T3, or T4) than the baseline utterance (T1). This procedure yielded measures corresponding to the percentage of trials that each listener detected convergence for each of the ten talkers producing each of the four phrases at each of the three time comparisons in the AXB test (except that one of the pairs did not provide utterances at T4). The data were submitted to a repeated measures analysis of variance to assess the effects of *Time Interval* (T2, T3, and T4), *Pair* (1–4), and *Phrase* (*dark suit*, *greasy*, *oily rag*, or *wash water*).

In comparison to speech produced at T1, listeners detected phonetic convergence in roommate pairs at T2 in October (55%), T3 in December (56%), and at T4 in January (56%). The main effect of *Time Interval* was not significant, indicating no differences in convergence across these intervals ( $F(2, 58)=1.345$ ,  $p=0.269$ ,  $\eta_p^2=0.044$ ). Despite the lack of difference across the intervals, the 95% confidence intervals from the analysis indicated greater than chance detection (50%) at all intervals. Each pair of talkers exhibited different levels of phonetic convergence, with all but one pair showing significant convergence overall (Pair 1=57%, Pair 2=52% ns, Pair 3=59%, Pair 4=55%;  $F(3, 87)=6.577$ ,  $p<0.001$ ,  $\eta_p^2=0.185$ ). Detection of phonetic convergence differed across the phrases, with greater levels of phonetic convergence on *wash water* (58%) and *oily rag* (57%) than on *dark suit* (54%) and *greasy* (53%;  $F(3, 87)=7.369$ ,  $p<0.001$ ,  $\eta_p^2=0.203$ ; 95% confidence intervals from the analysis indicated greater than chance detection for all phrases and confirmed the observed differences between phrases).

Examining the data more closely, the roommate pairs showed different patterns of phonetic convergence across the time intervals. As shown in Fig. 1, two of the pairs increased in convergence from T2 in October to T3 in December and decreased at T4 in January, when they had just returned from winter break. Pair 3 showed the opposite pattern, with the highest levels of convergence at T2 and T4 and a decrease at T3, and Pair 2 did not converge until T4. The interaction between *Time Interval* and *Pair* was significant ( $F(6, 174)=2.897$ ,  $p=0.01$ ,  $\eta_p^2=0.091$ ; 95% confidence intervals from the analysis indicated greater than chance detection for all pairs at all time intervals except T2 and T3 for Pair 2). Although not included in the statistical analyses, the convergence measures for Pair 5 showed a marked decline from T2 (58%) to T3 (46%).

Finally, each pair of talkers demonstrated a distinct pattern of phrase-dependent phonetic convergence. As shown in Fig. 2, listeners detected convergence on *oily rag* and *wash water* for Pair 1; *dark suit* and *wash water* for Pair 2; *dark suit*, *oily rag*, and *wash water* for Pair 3, and *greasy* and *oily rag* for Pair 4. The interaction between *Pair* and *Phrase* was significant ( $F(9, 261)=11.963$ ,

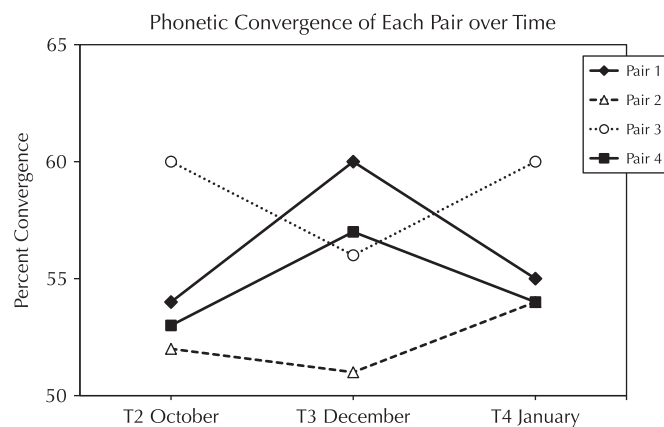


Fig. 1. Phonetic convergence of each roommate pair varies over the course of the academic year. Convergence is measured as percent detection of increased similarity at each time interval. All values differ from 50% chance detection, except for Pair 2 at T2 and T3.

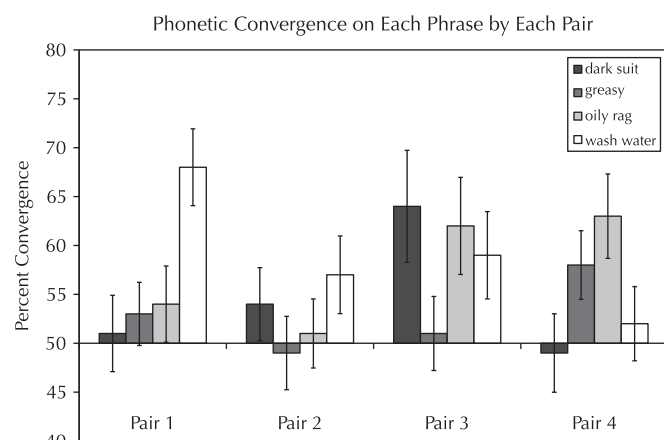


Fig. 2. Each roommate pair shows a distinct pattern of phonetic convergence across different phrases. Convergence is measured as percent detection of increased similarity averaged across time intervals. Error bars depict 95% confidence intervals.

$p<0.001$ ,  $\eta_p^2=0.292$ ; error bars depict 95% confidence intervals). The interactions between *Time Interval* & *Phrase* and *Time Interval*, *Pair*, & *Phrase* were not significant ( $p=0.115$ ,  $0.263$ ;  $\eta_p^2=0.056$ ,  $0.039$ ).

The location of origin and perceived convergence of each individual talker within a pair are presented in Table 1, averaged across time intervals. Six of the talkers in this study originated from locations in New York State, with other talkers from New Jersey, Tennessee, Florida, and Korea. The pairs with talkers from the most distinct regions were pairs 2, 3, and 4. Overall, pair 3 demonstrated the greatest and most evenly balanced levels of perceived convergence. Pairs 1 and 4 demonstrated the next highest levels of phonetic convergence, but their convergence was asymmetrical. (In the case of Pair 4, this asymmetry could be due to the fact that talker A was born in Korea, although he reported that he was a native English speaker, but Pair 1 showed a similar pattern.) Pairs 2 and 5 demonstrated the lowest levels of convergence overall. Therefore, phonetic convergence levels were not consistently related to distance in region of origin, as those pairs from the most distinct regions exhibited both the highest and lowest levels of convergence.

Because each talker pair converged on different phrases, it is unlikely that the measure of phonetic convergence simply reflected a shift toward sounding more like New Yorkers. For example, New Englanders typically produce *dark suit* without the



**Table 1**

Location of origin and convergence data for individuals participating in the recording sessions.

Talker A	Talker B	Perceived convergence of A to B	Perceived convergence of B to A	Vowel distance change <sup>a</sup>
1 Peekskill, NY	New York, NY	54	59	41
2 Syracuse, NY	Miami, FL	53	52	–18
3 Huntingdon, TN	New York, NY	59	59	–115
4 Seoul, Korea <sup>b</sup>	Pleasant Valley, NY	51	58	–14
5 Syosset, NY	Brick, NJ	53	51	37

Location of origin corresponds to the place where the individual lived for the longest period of time prior to participating in this experiment. Perceived convergence has been averaged across the time intervals.

<sup>a</sup> Vowel distance changes are in F1' by F2' Hz space. The Euclidean distances between paired talkers in the time 1 session were subtracted from paired talker distances in the subsequent sessions and then averaged. Negative values indicate a reduction in paired vowel distances from the first to the subsequent sessions.

<sup>b</sup> This individual lived in Fullerton, CA for 4 years prior to participating in this experiment.

post-vocalic *r* and *greasy* with a voiceless fricative (Clopper & Pisoni, 2004). However, in this group, all talkers produced *dark suit* with post-vocalic *r* consistently, so the perceived convergence on *dark suit* in pairs 2 and 4 and lack of convergence in the other pairs is most likely due to other phonetic attributes. This could be due to the fact that most talkers were from New York state, in which the *r*-less form would be found among members of lower socio-economic groups than are typically found attending Columbia University (Labov, 2006a). Furthermore, most of the talkers produced the voiceless version of *greasy* consistently, except for one talker from pair 2 and one from pair 4. Note that the only pair of talkers that was found to converge on *greasy* was pair 4. Because the listeners did not detect convergence in *greasy* for pair 2, but did detect convergence in *greasy* for pair 4 when both pairs differed in voicing, there must be other phonetic attributes that influenced perceptual similarity.

### 3.2. Duration analyses

In order to begin to identify potential acoustic–phonetic attributes that talkers might have converged on, the duration of the AXB items were analyzed. Moreover, it is possible that talkers were using more formal/careful speech on their first visit to the laboratory, and so the T1 items that were used as baselines would be distinct in duration from the items recorded at later time intervals. The item durations decreased on average from T1 to T2/T3/T4, indicating potential usage of more casual forms at later time intervals (T2: –16 ms; T3: –5 ms; T4: –29 ms;  $F(2, 32) = 5.54$ ,  $p = 0.009$ ,  $\eta_p^2 = 0.257$ ). However, the average differences are relatively small and there was variation in the direction and degree of difference in duration across items and pairs. Some items for some pairs at some time intervals were actually longer in duration than the T1 items.

In order to determine whether listeners were responding to average duration when making their perceptual similarity judgments, the differences in duration from T1 to T2/T3/T4 and the phonetic convergence data were submitted to a correlational analysis. The correlation between the perceived convergence of each item at each time interval in each pair and the difference in duration between T1 and the relevant counterpart at each time interval in each pair was not significant ( $r(46) = -0.08$ ,  $p > 0.05$ ). Therefore, despite the small but significant reduction in duration from T1 to later time periods, the pattern of phonetic convergence

as detected by the listeners was unrelated to the pattern of duration differences. This finding echoes the previously reported lack of relationship between duration and phonetic convergence for male talkers, and the failure to find a relationship between articulation rates and phonetic convergence (Pardo, 2010; Pardo et al., 2010).

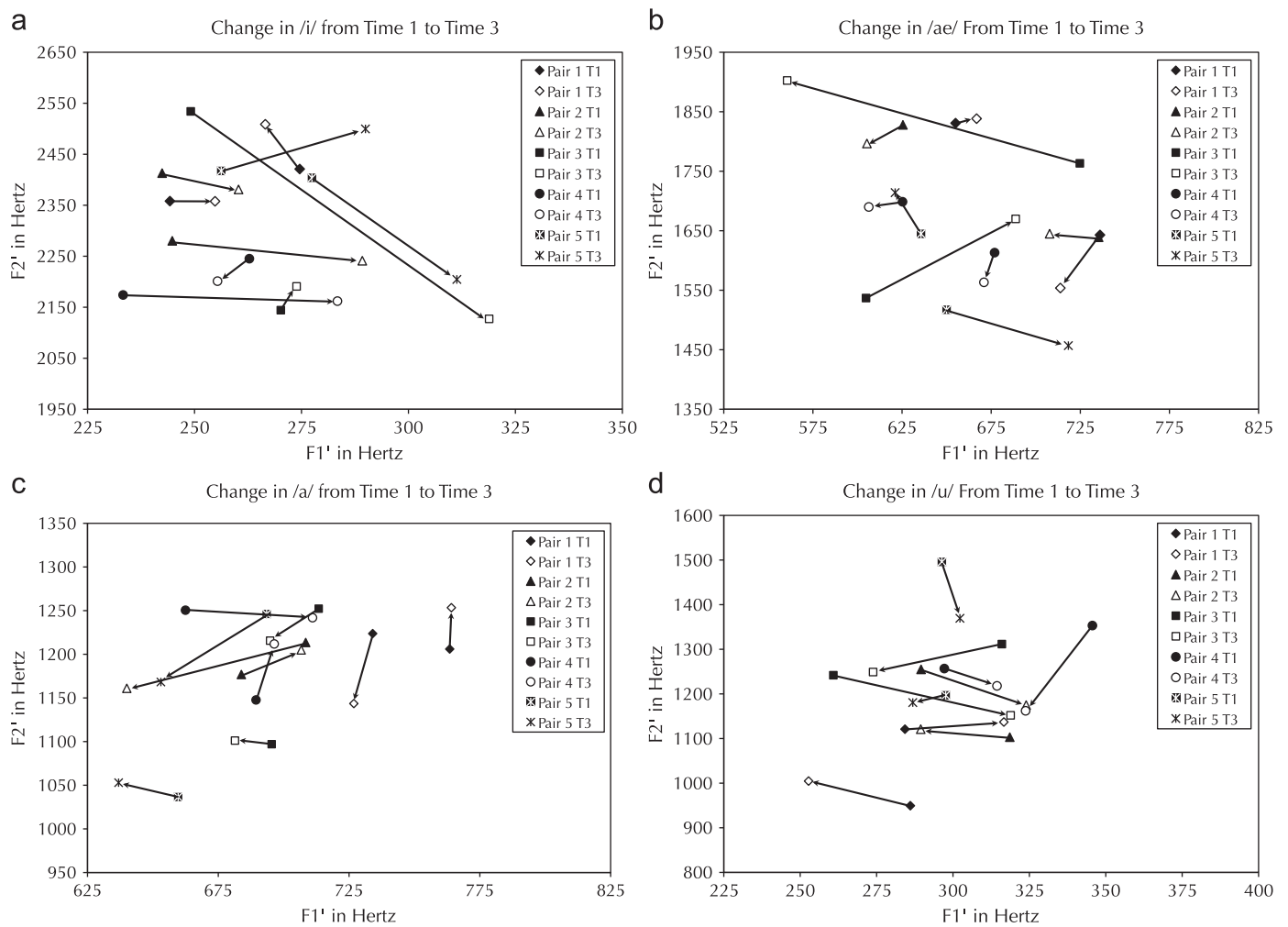
### 3.3. Vowel measures

In addition to the items used in the perceptual similarity tests, each talker also produced 5 repetitions of the full set of American English vowels embedded in hVt/d words in a carrier sentence. Because the items used to assess phonetic convergence also contained all of the point vowels (/i/, /æ/, /a/, /u/), the repetitions of the point vowels from the full set were analyzed for comparison with the perceptual measures. The averages of the first and second formants across the vocalic portion of each vowel token were derived using Praat estimates (Boersma & Weenink; [www.praat.org](http://www.praat.org)). Then, the estimates were normalized in order to reduce the impact of anatomical differences between talkers, yielding measures of F1' and F2'. The normalization routine preserves dialectal and ideolectal differences while projecting each talker's formant values into a common acoustic space (see Labov, 2006b, 2006c; Nearey, 1989). The data were scaled in a single batch using the North Carolina State University Linguistics Program's online utility with the Labov ANAE extrinsic setting (Thomas & Kendall, 2007; Thomas, Kendall, Yeager-Dror, & Kretzschmar, 2007).

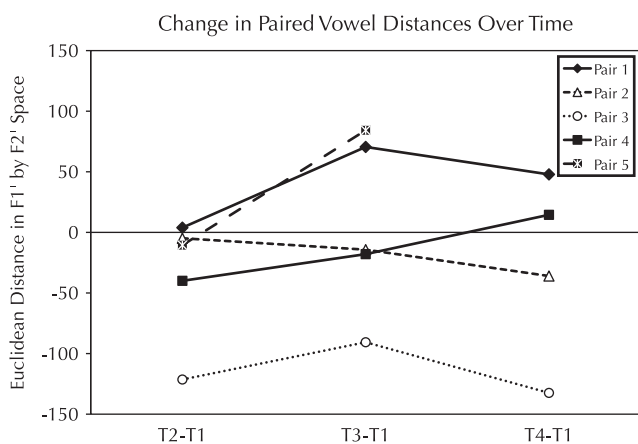
In order to simplify presentation of the vowel measures, the average F1' by F2' measures from T1 (August) and T3 (December) for each talker are plotted in four panels in Fig. 3. Each panel corresponds to a single vowel, /i/, /æ/, /a/, or /u/. The filled bullets depict averages from T1, and the open bullets depict averages from T3. Each talker pair is represented by the same bullet shape. Arrows connect the vowels of the same talker from T1 to T3.

Across all panels in Fig. 3, the change in vowel formants from T1 to T3 is extremely complex. The arrows do not all move in the same direction, nor do they converge toward a single point for any of the panels. In the panels depicting /i/ and /u/, some of the talkers appear to be moving toward the center of the vowel space (mainly increasing F1' and decreasing F2'), but others move in the opposite or different directions. If the talkers had all been converging toward a local New York City dialect, then the formants should have started to shift in a more uniform fashion for at least one of the vowels (Evans & Iverson, 2007). Unfortunately, the roommates also do not appear to be moving closer to each other in their vowel formants from T1 to T3. The only instances in which a pair of roommates' vowels moved closer together from T1 to T3 were pair 4's average /u/ and /a/ vowels.

In order to analyze the changes in vowel formants, each roommate pair's vowels were first converted to Euclidean distances. These paired distances were submitted to a mixed-design Analysis of Variance to test for the within-subjects effects of *Time Interval* and *Vowel*, and the between subjects effects of *Pair*. The Euclidean distance between the roommate pairs' vowels reduced from T1 (205) to T2 (171), then increased at T3 (212), and decreased again at T4 (189), and the pattern was significant ( $F(3, 60) = 5.394$ ,  $p < 0.05$ ;  $\eta_p^2 = 0.212$ ; 95% confidence intervals from the analysis indicated that T1 and T3 differed from T2 and T4). However, each roommate pair differed in their distances over time. Fig. 4 displays the differences in roommate pair distances from T1 to the later time intervals. Negative differences indicate a reduction in vowel distances over time. As the figure shows, pairs 2, 3, and 4 either reduced vowel distances or showed no overall change from T1 to later time intervals, and pairs 1 and 5 showed the opposite pattern. The interaction between *Time Interval* and *Pair* was significant ( $F(12, 60) = 4.595$ ,  $p < 0.05$ ;  $\eta_p^2 = 0.479$ ). The



**Fig. 3.** (a)–(d) Normalized vowel formants show inconsistent changes from T1 (August) to T3 (December). Each panel plots a single vowel. Averages at T1 are depicted with filled bullets, and averages at T3 are depicted with open bullets. Roommate pairs are plotted with bullets of the same shape. Arrows indicate direction of movement for each individual talker from T1 to T3.



**Fig. 4.** The Euclidean distances between roommate's vowels changed from T1 (August) to T2 (October), and only some pairs demonstrated additional changes at T3 (December) and T4 (January). Each bullet represents the difference in roommate vowel distances from T1 to T2, T3, or T4.

average differences in distances over time are shown in the right column of Table 1, alongside the perceived convergence for each talker in each pair. Although there appears to be a relationship between perceived convergence and average change in vowel

distances, there are not enough pairs in the current study to establish a reliable pattern ( $r(3) = -0.64$ , ns).

### 3.4. Relationship quality

The last set of analyses assesses the relationship between the roommates' perceived convergence and relationship quality. Roommate relationship quality was assessed at T3 using several survey questions, including number of waking hours spent in the same room per week, number of meals eaten together per week, seven-point scales measuring closeness and amount in common, and the IOS (Aron et al., 1992). Data on number of meals shared per week were excluded as each roommate indicated that the number was zero. There were significant correlations among the data for three of the measures, amount in common, closeness, and IOS, but number of hours/week spent together did not correlate with any of these measures. Therefore, a composite of the correlated relationship measures, amount in common, closeness, and IOS, was formed by averaging the ratings for the three measures, yielding one composite closeness measure for each individual.

In order to determine whether an individual's closeness to his roommate was related to his own convergence to his roommate, independent of whether his roommate shared his closeness sentiments or whether his roommate displayed convergence,

the estimates, ratings, and composite Closeness measures were correlated with the individual convergence measures from each of the 10 talkers at T3 in December. This analysis revealed a significant correlation between rated closeness and convergence at T3 ( $r(8)=0.54$ ,  $p=0.05$ ) and a modest correlation between the composite Closeness index (closeness, amount in common, and IOS rating) and convergence at T3 ( $r(8)=0.36$ ,  $p=0.15$ ). No other correlations with the convergence data at T3 were significant (nor did the relationship quality measures taken at T3 correlate with perceived convergence at any of the other time intervals). These findings suggest that a talker's degree of phonetic convergence to a roommate after approximately 3.5 months of cohabitation is related to positive levels of reported closeness to his roommate at that time.

#### 4. Discussion

This study examined phonetic convergence in five pairs of previously unacquainted college roommates at three time intervals during the academic year. Each member of each pair provided multiple samples of phrases that were submitted to AXB perceptual listening tests. In addition, each roommate provided measures of their perceived closeness, collected in December after the roommates had been cohabitating for approximately 3.5 months. As suggested by previous research, most roommates converged in perceived phonetic form within the first time interval after approximately 1.5 months of cohabitation, perceived phonetic convergence was evident for the same pairs at the second time interval after approximately 3.5 months, and all roommates converged by the last time interval, after returning from winter break. Measures of item duration differences and vowel spectra were not related to perceived convergence. The degree and patterns of convergence varied across pairs, phrases, and measures.

These findings are compatible with proposals that follow from Communication Accommodation Theory (Shepard et al., 2001). For example, Gallois et al. (1995) suggested that convergence may result from a desire to make an interaction flow more smoothly, a desire that can be reasonably attributed to the roommates in this study. Moreover, studies of social interaction have shown that other forms of behavioral mimicry lead to impressions of smoother interactions and greater liking for a partner (Chartrand & Bargh, 1999; Chartrand, Maddux, & Lakin, 2005; Lakin & Chartrand, 2003). Therefore, it is likely that phonetic convergence of the kind observed in individual conversational interactions and in the current study reflects a desire to decrease social distance and to induce a smooth interaction and mutual liking. Additional research is necessary to detail the manner in which such factors interface with speech perception and production to evoke convergence in phonetic form.

The current study found variable patterns of phonetic convergence over the course of the academic year, both across different pairs and different utterances. Although previous findings based on a single talker who moved between two different linguistic environments (Sancier & Fowler, 1997) suggested that convergence should be reduced after winter break, only two of the five pairs showed this pattern, and one pair only showed significant convergence after returning from winter break. One important difference between these two studies was the methods that were employed. The perceptual similarity test used in the current study is a sensitive measure of paired talker phonetic similarity, whereas an accentedness rating task likely focuses on the impact of the intonation pattern of a second language environment on a talker's native language. The acoustic attribute that was measured by Sancier and Fowler (VOT) was selected because it is a

phonological variant that is known to vary in its distribution between the two languages. However, VOT is also correlated with speaking rate, a relatively coarse-grain attribute (Miller & Grosjean, 1981). Likewise, Evans and Iverson (2007) reported changes in accentedness ratings and vowel formants for college students in England. Unfortunately, the students in the current study did not demonstrate similarly consistent shifts in vowel formants, which is probably due to the fact that most of them were from the New York state area and had already been attending Columbia University for at least one year. They all produced *r*-full utterances of *dark suit*, and variability in fricative voicing in *greasy* was not related to perceived convergence.

The current results suggest that adjustments in perceived phonetic repertoire follow a trajectory that differs from that of accentedness or VOT patterns, appearing more resistant to decay across breaks in exposure. However, this could be due to the fact that the talker in the study by Sancier and Fowler was alternating between different languages with large phonetic differences. In order to determine whether within-language phonetic convergence differs from gestural drift between languages, it will be necessary to manipulate the range of variation that a talker explores within their own language.

Previous studies of communication accommodation focused on attributes that were measured in much longer time-scales, including accent, speaking rate, intensity, low-frequency band variation, pause frequency, and utterance length (Giles et al., 1991; Gregory, 1990; Gregory et al., 1997; Gregory & Webster, 1996; Natale, 1975). In most of these reports, paired talkers were found to converge in the particular attribute that was measured. However, Pardo et al. (2010) failed to find evidence of convergence in speaking rate at the same time that talkers converged phonetically, and roommates in the current study did not converge on vowel spectra consistently. In a now classic study by Bilous and Krauss (1988), talkers were found to converge in some attributes at the same time that they diverged in others (e.g., average utterance length; frequency of pauses, laughter, interruptions, and back-channels).

These findings lead to the intriguing possibility that each individual talker might converge on a unique set of acoustic-phonetic attributes while diverging, varying randomly, or remaining neutral on others. If that is the case, then measurements of individual acoustic attributes will yield inconsistent patterns. Therefore, a complete understanding of phonetic convergence is unlikely to result from acoustic analyses alone. Without a perceptual similarity assessment, a failure to find convergence in a particular acoustic-phonetic attribute cannot be interpreted. Moreover, a finding of increased similarity in any acoustic-phonetic attribute must be interpreted against the background of a talker's complete phonetic repertoire. At this point, the most effective way to assess these relationships is to rely on the judgments of ordinary listeners. Because ordinary perception integrates multiple dimensions simultaneously, a carefully designed perceptual similarity test provides a global assessment of phonetic convergence without committing to a single acoustic-phonetic attribute that might not be used consistently by every talker on every occasion.

It is important to point out that the overall levels of detected convergence in college roommates was modest, even after 3.5 months of relatively continuous cohabitation. These findings align with those found in studies of convergence in shadowing tasks and during conversational interaction (Babel, 2010; Goldinger, 1998; Namy et al., 2002; Pardo, 2006; Pardo et al., 2010; Shockley et al., 2004). If phonetic convergence is automatically evoked by perceptual resolution of phonetic forms that goad imitative production (Fowler et al., 2003; Pickering & Garrod, 2004), or by automatic activation of relatively recent



episodic traces (Goldinger, 1998), then any two talkers who live together should exhibit phonetic convergence with a high degree of fidelity, and the level of convergence should not vary across different utterances. The fact that talkers never match acoustic-phonetic attributes exactly despite a putative drive toward parity calls into question accounts that rely on a direct perception-production link in determining the phonetic form of utterances (Fowler et al., 2003; Pickering & Garrod, 2004). These findings indicate that phonetic convergence in naturalistic settings is not an automatic consequence of a direct perception-production link (Pardo et al., 2010), and that the social and situational factors at play evoke variability in phonetic form that is convergent, neutral, and divergent.

Despite its limitations, the current study offers a useful paradigm for examining phonetic variation that results from natural social interaction. To varying degrees, talkers have been found to imitate a model in shadowing tasks, to converge toward a conversational partner during a single interaction, and to show longer-term adjustments in phonetic repertoire as a result of continued contact with the same talker. Additional research is needed to extend the current findings both to female pairs and to mixed sex cohabitation settings. Ultimately, a complete account of variability in phonetic repertoire across both individuals and groups will examine not only individual pairs, but communities of talkers interacting across multiple settings.

## Acknowledgments

Completion of this paper was supported in part by Grant #0545133 from the National Science Foundation to Jennifer Pardo at Barnard College. The authors are indebted to Robert Remez, Isabel Cajori Jay, and the reviewers for their role in the completion of this project.

## References

- Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, 63, 596–612.
- Babel, M. (2009). *Phonetic and social selectivity in speech accommodation*. Doctoral Dissertation. Berkeley: University of California.
- Babel, M. (2010). Dialect convergence and divergence in New Zealand English. *Language in Society*, 39, 437–456.
- 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.
- Byrne, D. (1971). *The attraction paradigm*. New York: Academic 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. W., & Lakin, J. L. (2005). Beyond the perception-behavior link: The ubiquitous utility and motivational moderators of non-conscious mimicry. In: R. R. Hassin, J. S. Uleman, & J. A. Bargh (Eds.), *The new unconscious* (pp. 334–361). New York, NY: Oxford University Press.
- Clopper, C. G., & Pisoni, D. B. (2004). Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics*, 32, 111–140.
- Eckert, P. (1989). *Jocks and burnouts: Social categories and identity in the high school*. New York: Teachers College Press.
- Evans, B. G., & Iverson, P. (2007). Plasticity in vowel perception and production: A study of accent change in young adults. *Journal of the Acoustical Society of America*, 121, 3814–3826.
- 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.
- Giles, H., Coupland, J., & Coupland, N. (1991). *Contexts of accommodation: Developments in applied sociolinguistics*. Cambridge: Cambridge University Press.
- Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105(2), 251–279.
- 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 relation of vocal accommodation in conversational partners' fundamental frequencies to perceptions of communication quality. *Journal of Nonverbal Behavior*, 21, 23–43.
- Gregory, S. W., & 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.
- Labov, W. (1972). The recent history of some dialect markers on the island of Martha's Vineyard, Mass. In: L. M. Davis (Ed.), *Studies in linguistics in honor of Raven I. McDavid Jr.* Alabama: University of Alabama Press.
- 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). New Jersey: Lawrence Erlbaum Associates.
- Labov, W. (2006a). *The social stratification of English in New York City* (2nd ed.). Cambridge: Cambridge University Press.
- Labov, W. (2006b). *The Atlas of North American English*. New York: Mouton.
- Labov, W. (2006c). A sociolinguistic perspective on sociophonetic research. *Journal of Phonetics*, 34, 500–515.
- Lakin, J. L., & Chartrand, T. L. (2003). Using unconscious behavioral mimicry to create affiliation and rapport. *Psychological Science*, 14, 334–339.
- Miller, J. L., & Grosjean, F. (1981). How the components of speaking rate influence the perception of phonetic segments. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 208–215.
- 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. (1975). Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality & Social Psychology*, 32, 790–804.
- Nearey, T. M. (1989). Static, dynamic, and relational properties in vowel perception. *Journal of the Acoustical Society of America*, 85, 2088–2113.
- Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America*, 119(4), 2382–2392.
- Pardo, J. S. (2010). Expressing oneself in conversational interaction. To appear. In: E. Morsella (Ed.), *Expressing oneself/expressing one's self* (pp. 183–196). Taylor & Francis.
- Pardo, J. S., Cajori Jay, I., & Krauss, R. M. (2010). Conversational role influences speech imitation. *Attention, Perception, & Psychophysics*, 72, 2254–2264.
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral & Brain Sciences*, 27, 169–190.
- 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*. Chichester, UK: John Wiley & Sons, Ltd.
- Shockley, K., Sabadini, L., & Fowler, C. A. (2004). Imitation in shadowing words. *Perception & Psychophysics*, 66(3), 422–429.
- Street, R. L., & Giles, H. (1982). Speech accommodation theory: A social cognitive approach to language and speech behavior. In: M. Roloff, & C. Berger (Eds.), *Social cognition and communication*. Beverly Hills, CA: Sage.
- Thomas, E. R., & Kendall, T. (2007). NORM: The vowel normalization and plotting suite. [Online Resource: <http://ncslaap.lib.ncsu.edu/tools/norm/>].
- Thomas, E. R., Kendall, T., Yeager-Dror, M., & Kretzschmar, W. (2007). Two things sociolinguists should know: Software packages for vowel normalization, and accessing linguistic atlas data. In *Proceedings of the workshop at new ways of analyzing variation (NWAV)* (Vol. 36). Pennsylvania, PA: University of Pennsylvania.
- Triandis, H. C., & Triandis, L. M. (1960). Race, social class, religion, and nationality as determinants of social distance. *Journal of Abnormal and Social Psychology*, 61, 110–118.
- 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.