PLEASE RETURN YOUR CORRECTIONS WITHIN 72 HOURS

JSLHR

Dear Author(s):

Please return your changes, comments, and/or your page proofs within 72 hours to:

Kathleen Halverson
Production Editor, *JSLHR*
American Speech-Language-Hearing Association
2200 Research Boulevard #250
Rockville, MD 20850-3289
Phone: 301-296-8719
Fax: 301-296-8587
khalverson@asha.org

To make corrections or comments:
1. Send an e-mail or e-mail attachment listing your changes or comments (preferred method)

or

2. Mark up the page proofs and return them in one of two ways: (a) by fax (please be sure to print clearly and leave enough room in the margin so that your changes are not cut off during transmission) or (b) by overnight mail to the address listed above.

Please make sure any author queries, listed on the last page of the file, are answered and that any special characters (such as phonetic symbols) are correct.

To order reprints:
Ask the production editor to provide you with prices and shipping information.

JSLHR

Article

The Influence of Auditory Acuity on Acoustic Variability and the Use of Motor Equivalence During Adaptation to a Perturbation

Jana Brunner,[a] Satrajit Ghosh,[a] Philip Hoole,[b] Melanie Matthies,[c]
Mark Tiede,[a,d] and Joseph Perkell[a]

Purpose: The aim of this study was to relate speakers' auditory acuity for the sibilant contrast, their use of motor-equivalent trading relationships in producing the sibilant /ʃ/, and their produced acoustic distance between the sibilants /s/ and /ʃ/. Specifically, the study tested the hypotheses that during adaptation to a perturbation of vocal-tract shape, high-acuity speakers use motor equivalence strategies to a greater extent than low-acuity speakers in order to reach their smaller phonemic goal regions, and that high-acuity speakers produce greater acoustic distance between two sibilant phonemes than low-acuity speakers.

Method: Articulographic data from 7 German speakers adapting to a perturbation were analyzed for the use of motor equivalence. The speakers' produced acoustic distance between /s/ and /ʃ/ was calculated. Auditory acuity was assessed for the same speakers.

Results: High-acuity speakers used motor equivalence to a greater extent when adapting to a perturbation than did low-acuity speakers. Additionally, high-acuity speakers produced greater acoustic contrasts than low-acuity-speakers. It was observed that speech rate had an influence on the use of motor equivalence: Slow speakers used motor equivalence to a lesser degree than fast speakers.

Conclusion: These results provide support for the mutual interdependence of speech perception and production.

Key Words: articulation, palate, speech sound, speech intelligibility

A number of experimental observations have suggested that speech perception and speech production are linked. Newman (2003), for example, found that subjects who had longer voice onset times (VOTs) in their prototypical /pɑ/ percepts also produced longer VOTs. Bradlow, Pisoni, Akahane-Yamada, and Tohkura (1997) investigated the /r/–/l/ contrast of Japanese learners of English during various training sessions and found that the speakers with a better perceptual distinction also produced a clearer contrast. Following from these observations, one can hypothesize that there are differences in speakers' auditory acuity: Some speakers are able to hear more subtle differences between speech sounds than others. As a consequence, high-acuity speakers will accept a smaller range of sounds as good representatives of a phoneme, whereas low-acuity speakers will accept a larger range of sounds. In speech production, high-acuity speakers will try to reach smaller phonemic goal regions and thus produce clearer contrasts between phonemes than low-acuity speakers. This hypothesis is consistent with the *Directions Into Velocities of Articulators (DIVA)* model of speech production (Guenther, 1994, 1995; Guenther, Ghosh, & Tourville, 2006; Guenther, Hampson, & Johnson, 1998), which assumes that speakers have auditory goal regions for each phoneme that vary in size and separation across speakers, and that speakers try to reach these goal regions in speech production.

Perkell, Guenther, et al. (2004) assessed speakers' auditory acuity and their produced contrast for the vowel

---

[a]Speech Communication Group, Massachusetts Institute of Technology, Cambridge, MA
[b]Institut für Phonetik und Sprachverarbeitung, Ludwig-Maximilians-Universität München, Munich, Germany
[c]Department of Speech, Language and Hearing Sciences, Boston University, MA
[d]Haskins Laboratories, New Haven, CT

Correspondence to Jana Brunner: jana.brunner@uni-potsdam.de

Jana Brunner is now with the German Department, University of Potsdam, Potsdam, Germany.

Editor: Robert Schlauch
Associate Editor: Julie Liss

Received December 1, 2009
Revision received July 1, 2010
Accepted September 21, 2010
DOI: 10.1044/1092-4388(2010/09-0256)

pairs /ɑ/–/ʌ/ and /u/–/ʊ/ and found cross-subject correlations between acuity and articulatory contrast and between acuity and acoustic phonemic contrast. Acuity was measured via a discrimination test (ABX) on seven-step *cod–cud* and *who'd–hood* continua. Subjects who discriminated all productions differing by just two steps on a continuum correctly were categorized as "high discriminators" for that continuum. The other subjects were categorized as "low discriminators." The same subjects' articulatory contrast for each vowel pair was measured from electromagnetic midsagittal articulography (EMMA) data as the Euclidean distance between the centroids of distributions of locations of a point on the tongue dorsum in the midsagittal plane from multiple repetitions of the vowels in normal, clear, and fast speech. Analogously, acoustic contrast distance was calculated as the Euclidean distance between centroids of distributions in first formant (F1) × second formant (F2) space. The results showed that the high discriminators produced greater acoustic and articulatory contrast distances than the low discriminators. In fast speech, both articulatory contrast and acoustic contrast were lower for *who'd* versus *hood* but not for *cod* versus *cud*.

Perkell, Matthies, et al. (2004) investigated the relation between auditory acuity and the stability of a particular articulatory pattern in producing the /s/–/ʃ/ contrast. Twenty speakers were recorded pronouncing the words *said, shed, sod,* and *shod* in normal, clear, and fast speech. Production of the sibilants /s/ and /ʃ/ can be distinguished in principle by the absence of a sublingual cavity for /s/ versus the presence of a sublingual cavity for /ʃ/. In order to investigate how consistently speakers produced this difference, a sensor that registered tongue contact with the lower alveolar ridge was placed on the alveolar ridge below the lower incisors. If contact was registered between the tongue and lower alveolar ridge, it was assumed that there was no sublingual cavity. The results showed that some speakers produced a very clear articulatory contrast (/s/ always without a sublingual cavity and /ʃ/ always with a sublingual cavity), whereas others produced either a less-clear or no-contact distinction between the two sibilants. Speakers were correspondingly grouped according to their "high" and "low" use of differential contact in producing the sibilants. The acoustic contrast between /s/ and /ʃ/ was calculated as the difference between the spectral means of the speakers' midsibilant productions. In addition, their auditory acuity was assessed using two synthesized *said–shed* continua, one with a male voice and one with a female voice. Following Perkell, Guenther, et al. (2004), speakers were at first asked to label each token as either *shed* or *said*. Afterwards, an ABX discrimination test was carried out in which speakers were asked to choose whether the third item (X) repeated the first or the second item. The results showed that the speakers with a clear articulatory contrast also had a greater acoustic contrast. Similarly, high-acuity speakers also had a clearer acoustic contrast than low-acuity speakers. In fast speech, all speakers tended to produce smaller acoustic contrasts.

Ghosh et al. (2010) measured auditory acuity of 18 speakers and related it to somatosensory acuity of the tongue tip and the magnitude of the sibilant acoustic contrast. Somatosensory acuity was assessed by testing how well speakers were able to perceive the orientation of grooves of different widths on a plastic dome pressed against the tongue tip. In order to investigate the acoustic sibilant contrast for the different speakers, sibilant spectral mean, skewness, and kurtosis were measured from the speakers' productions of *said, shed, sid,* and *shid*. The speakers' auditory acuity was also measured using an adaptive staircase just-noticeable-difference (JND) task. Contrast distance was measured as the average Euclidean distance between /s/ and /ʃ/ in three-dimensional (3-D) space defined by the three spectral moments. The results showed that the subjects' produced acoustic contrast distance between the sibilants was related both to their somatosensory acuity and their auditory acuity.

In the present study, we further investigated the role of auditory acuity in speech. One of the aims of the present study was to investigate a possible link between auditory acuity and an articulatory strategy—that is, the extent to which speakers use motor equivalence when they adapt to a perturbation.

Motor Equivalence

The term *motor equivalence* is used to denote different articulatory configurations that result in the same or a very similar acoustic output. Bite-block studies, for example, have shown that, if a speaker's jaw movement is blocked, the speaker can use the tongue to compensate and produce nearly normal speech by forming the same constriction sizes and locations as in unperturbed speech (Kelso & Tuller, 1983; Lindblom, Lubker, & Gay, 1979). Some studies, however, have found that there are small acoustic differences between the bite-block and the normal condition in vowels and consonants (e.g., McFarland & Baum, 1995).

Other studies have demonstrated motor equivalence by showing that speakers produce the same vocal-tract shape with different contributions of the articulators when lip movements are perturbed (Folkins & Zimmermann, 1982; Gracco & Abbs, 1985). Folkins and Zimmermann (1982) carried out a perturbation experiment in which the speaker's lower lip was moved down unexpectedly by electrical stimulation of the depressor labii inferior. The three speakers in the study showed active compensatory behavior in bilabial stop production by moving the upper lip further down and the jaw upward.

Gracco and Abbs (1985) perturbed the movement of the lower lip by pushing it down unexpectedly with a paddle during bilabial stop production. Speakers compensated via an increase in movement amplitude, velocity, and movement time of both upper and lower lips in order to produce the intended closure. In these studies, speakers were able to produce the acoustically critical constriction with varying contributions of the contributing articulators (jaw, lips, and tongue).

Different articulatory configurations that result in similar acoustic outputs have been found in the vowel /u/, American English /r/, and, in a few cases, /ʃ/. Following the nomograms in Fant (1960), the sound /u/ can be produced with protruded lips and a constriction in the velar region, but also with open lips and a constriction in the velopharyngeal region. Perkell, Matthies, Svirsky, and Jordan (1993) investigated motor-equivalent trading relations between lip rounding and rearward tongue-body raising in speakers' multiple repetitions of American English /u/. Both movements contribute to lowering F2 and could, therefore, covary while keeping the perceptually important acoustic cue relatively constant. Perkell et al. (1993) analyzed EMMA measurements of productions of about 300 repetitions of /u/ in different contexts by each of four speakers. Three of the four subjects showed weak negative correlations between the two parameters, which is consistent with the motor equivalence hypothesis.

Savariaux and colleagues (Savariaux, Perrier, & Orliaguet, 1995; Savariaux, Perrier, Orliaguet, & Schwartz, 1999) also investigated motor equivalence strategies for /u/. In these studies, speakers' lip movement was perturbed by a 2.5-cm diameter tube that held the lips open. Speakers were asked to produce /u/ in this condition. X-ray data recorded from the subjects showed that the majority of them retracted the tongue, resulting in a produced sound with formant frequencies that were similar to productions with protruded lips and a velar constriction.

American English /r/ can also be produced in at least two different ways—namely, with a bunched tongue or a retroflexed tongue (Delattre & Freeman, 1968). In both cases, a constriction is created that leads to a low third formant (F3), a very salient acoustic characteristic of this sound. Westbury, Hashi, and Lindstrom (1998) found for a sample of 53 speakers that the two production types—bunched and retroflexed—represent the endpoints of a continuum, and productions are often somewhere between the two extremes. The production configuration furthermore depends not only on the speaker but also on the context. Guenther et al. (1997) investigated /r/ production in seven speakers with EMMA recordings of the positions of points on the tongue. They found systematic articulatory tradeoffs, which reduced acoustic variability across contexts despite large variations in vocal-tract shape. The speakers covaried the length of the front cavity and the length and/or area of the palatal constriction, and both configurations resulted in a lowering of F3. It should be noted that small acoustic differences between the two production types have been found by Zhou et al. (2008). The retroflex variant has a larger difference between fourth formant (F4) and fifth formant (F5) than does the bunched variant. The authors explain this result by pointing out that the resonances are associated with different cavities in the two production types. This finding shows that the vocal-tract shapes of the two sounds indeed differ, although the most salient acoustic cue—a low value of F3—is very similar.

Motor equivalence in the production of /ʃ/ involves covariation of lip protrusion and position of the constriction formed by the tongue blade against the anterior palate. Speakers may produce this sound with protruded lips and an advanced constriction or without lip protrusion and a more retracted constriction. The effect of this covariation is that the size of the front cavity (lip opening to constriction) is kept approximately the same. Perkell et al. (2000) reported a study in which motor equivalence for lip rounding versus tongue-tip fronting was investigated in /ʃ/ for eight speakers. The results were mixed in that some speakers showed motor equivalence but others did not.

Motor Equivalence May Be a Means of Reducing Articulatory Effort

Some motor equivalence studies have perturbed an articulator and shown that another articulator compensates for it. Other studies have shown that speakers use different motor equivalence strategies in different contexts. When not perturbing speakers' speech and when not looking at different contexts, however, motor equivalence cannot be readily observed. This may be because motor equivalence can be understood as a means of reducing articulatory effort: Speakers use an articulatory strategy that is economic for the particular situation and at the same time leads to a sufficiently large acoustic contrast between neighboring sounds. If the context does not change and there is no perturbation, speakers will prefer their default strategy. Therefore, motor equivalence strategies can be shown only in the form of articulatory trading relations over many repetitions of the same sound (e.g., /u/), and in such examples, the evidence (negative correlation between two contributing articulatory movements) is relatively weak (Perkell et al., 1993).

Consider, as an example that is in line with the assumption that motor equivalence is a means of reducing articulatory effort, Westbury et al. (1998), who found that bunched /r/ is preferred in words such as *across* and *street,* whereas the retroflex variant is preferred, for example, in *row*. The choice of the strategy here can be explained by differences in articulatory effort for each variant in the different contexts: In *across,* the tongue forms a velar constriction with an already high tongue back for the velar stop. It should then involve less articulatory effort to move the tongue dorsum up for the bunched variant than to move the tongue back down and the tongue tip up for the retroflexed variant. In the word *street,* the choice of the articulatory strategy might be determined by competing demands on the articulators due to production context (coarticulation). While the tongue tip is still involved in the alveolar closure of /t/, the uninvolved tongue dorsum can begin to move upward for the formation of the bunched /r/. Waiting for the stop to be released and then quickly retroflexing the tongue tip in order to produce a retroflex /r/ in this same context would presumably require greater articulatory effort and would reduce the effects of coarticulation.

So, motor equivalence can also be a way to coarticulate efficiently, which might not seem plausible at first consideration: Whereas motor equivalence is a means to reduce articulatory effort and keep the acoustic output stable, coarticulation is also a means to reduce articulatory effort but often with the effect of inducing variability. However, as seen in the /r/ example from Westbury et al. (1998), these two phenomena do not compete with each other but can complement each other: In *street,* the only way to coarticulate efficiently is to use bunched /r/. If the retroflexed variant were to be used, /t/ and /r/ would have to be produced one after the other without the degree of articulatory overlap licensed by the bunched strategy.

As another example in which motor equivalence facilitates coarticulation, it has been found that there is more lip rounding in consonants surrounded by rounded vowels than in consonants surrounded by unrounded vowels (Benguerel & Cowan, 1974; Sussman & Westbury, 1981). A way to reduce the acoustic change induced by this coarticulation would be to use a motor equivalence strategy. In the case of /ʃ/, speakers could place the tongue in a slightly more advanced position in rounded context. Thus, motor equivalence can be a means of facilitating coarticulation while keeping the acoustic output relatively stable.

Related to this discussion is the notion of a *natural perturbation*. Edwards (1985) investigated the contribution of movements of the jaw and tongue on the formation of a /t/ closure in different contexts. She found that there was less variability in the combined tongue-and-jaw movement than in each single articulator movement (jaw movement and intrinsic tongue movement). Similar to a perturbation paradigm in which the jaw movement is blocked with a bite block, in Edwards's data the jaw might be seen as being "blocked naturally" by coarticulatory demands so that the tongue has to compensate for that (similarly Fowler & Saltzman [1993] for lip perturbation). Applied to /ʃ/ in rounded contexts, if speakers produce lip rounding because of the context and place the tongue at a more advanced position, the tongue compensates for the lip that is "blocked" by the context. These examples show that coarticulation and motor equivalence can be closely linked.

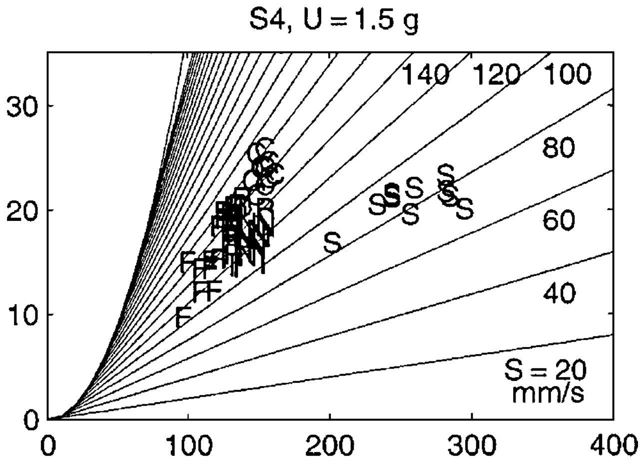
Reduction of Articulatory Effort May Be More Vital in Fast Than in Slow Speech

So far, we have argued that motor equivalence is important for efficient articulation, including coarticulation. We suggest further that reduction of articulatory effort (and, thus, motor equivalence) should be more relevant in fast speech than in slow speech.

This inference is based on results presented in Perkell, Zandipour, Matthies, and Lane (2002). Following Nelson (1983), Perkell et al. (2002) showed that movement time is not linearly related to the effort involved in carrying out a movement, quantified indirectly as *peak velocity* in the study by Perkell et al. (2002). The results for one of the speakers from Perkell et al. (2002) are given in Figure 1. The "F" symbols in this figure denote measures of tongue-blade opening movements for the utterance *tot* (in a carrier phrase) from a fast-speaking condition; the "S" symbols denote measures from a slow speech condition (for more detail, see Perkell et al., 2002). The abscissa shows duration in ms, and the ordinate shows movement distance. Effort is presented as radiating

FIG1

Figure 1. Results for one of the speakers from Perkell et al. (2002; see Figure 9 of Perkell et al. [2002], Speaker S4). Abscissa gives movement duration in ms; ordinate gives movement distance in mm. Numbers on the right give speed (a measurement of effort) in mm/s.



curved lines in the figure. The values for effort are given at the end of the lines (peak movement speed in mm/s).

For a given movement distance, the effort is low for long movement durations and does not increase much for a decrease in movement duration. However, for short movement durations, the effort increases dramatically even for a small reduction in movement time. Thus, if a speaker speaking slowly finds a way to reduce the articulatory effort (e.g., by using a motor equivalence strategy reducing movement distance), this does not have as great an effect on effort reduction as it would have if the speaker spoke faster.

Looking at the measurement results, the movements from slow productions all have longer durations and somewhat greater movement distances. The effort is much lower for the slow condition (about 80 mm/s) than for the fast condition (140 mm/s and beyond). Reducing the movement distance (as could be done by using a motor equivalence strategy) would have a greater effect on the effort in the fast than in the slow condition.

As a consequence of the two assumptions—that is, that motor equivalence is a means of reducing articulatory effort and that reduction of articulatory effort is more important in fast than in slow speech—there should be an observed tradeoff between motor equivalence and speech rate: In fast speech, speakers should use motor equivalence strategies to a higher degree than in slow speech.

Subject-Specific Differences in Motor Equivalence and Aim of the Study

The motor equivalence studies discussed earlier provide evidence of speakers' ability to use more than one kind of motor equivalence strategy: Speakers can produce the same vocal-tract shape by varying contributions of different articulators, and they can also use very different vocal-tract shapes that result in the same acoustic output, at least for some sounds. In many studies, however, a great deal of interspeaker variability was observed with respect to the extent to which motor equivalent behavior was observed. For example, in Folkins and Zimmerman (1982), patterns of motor equivalence were not as clearly shown by Subject 2 as they were by the other two subjects in the study. In the Perkell et al. (1993) study, the relation between tongue-body raising and lip protrusion was quite clear for three subjects, but one subject did not show motor equivalence. In the study by Savariaux et al. (1995), four of the 11 speakers did not adapt to the lip tube; their tongue shape and position were the same as without the lip tube. Perkell et al. (2000) observed a similar result: Not all of the subjects had motor equivalence patterns in /ʃ/ production. The aim of the present study was to find out why some speakers use motor equivalence and others do not. In line with findings suggesting that acuity influences articulation, the main hypothesis of the investigation was that speakers' use of motor equivalent trading relations could also be related to their acuity. The aim of motor equivalence strategies is to keep the acoustic output within an acceptable goal region. High-acuity speakers should thus use these strategies to a greater degree than low-acuity speakers because high-acuity speakers hypothetically have smaller goal regions than do low-acuity speakers. Furthermore, in line with previous findings (Perkell, Guenther, et al., 2004; Perkell, Matthies, et al., 2004), high-acuity speakers should produce greater (clearer) acoustic contrasts between phonemes than should low-acuity speakers. In order to investigate these relations, the acuity of the six speakers for the /s/–/ʃ/ contrast from Brunner and Hoole (2009) and one additional subject was measured and related to their patterns of motor equivalence. Additionally, the acoustic contrast distance of their /s/–/ʃ/ productions was measured and related to both auditory acuity and patterns of motor equivalence.

Method

Subjects

The subjects were all paid volunteers: three men (AM1, AM2, CM1) and four women (AF1, CF1, CF2, CF3). They were between 25 and 56 years of age. All speakers had no history of speech or hearing problems except for speaker CM1, who reported having had a slight left-sided hearing problem in the past that was not found to have any effect on his speech or hearing at the time of recording. All speakers spoke Standard German with some regional influences.

Experiment 1: Motor Equivalence and Acoustic Distance

In order to measure motor equivalence in the production of /ʃ/, a 2-week long perturbation experiment was carried out. The experiment is described in detail in Brunner and Hoole (2009). The aim of this previous study was to investigate motor equivalence in /ʃ/ for different palate shapes and under different feedback conditions. The motor equivalence measurements for six of the seven speakers (all except CM1) are taken from this earlier study. Relevant details on the methodology of the present study are repeated here.

Perturbation. The perturbation device was an artificial palate. There were two types of artificial palates—one that effectively displaced the alveolar ridge in a posterior direction ("alveolar palate"') and another that made the palate flatter and lower by filling in the palatal arch (*central palate*). Two different artificial palates were chosen for the original study because it was hypothesized that the speakers with an alveolar palate would use the alveolar ridge as a landmark during adaptation, so that the articulatory variability within the perturbed sessions would be reduced. The speakers with a central palate were expected to have more variability under perturbation because less information about the constriction location could be gained from tactile feedback. This, however, was not observed in Brunner and Hoole (2009).

Three speakers (AM1, AM2, and AF1) were recorded with an alveolar prosthesis, and four speakers (CM1, CF1, CF2, and CF3) were recorded with a central prosthesis. Speakers were asked to wear the prosthesis all day for 2 weeks and to make a serious effort to adapt to it in their speech.

Setup of experiment. The speakers were recorded via electromagnetic articulography (EMA) with sensors placed midsagittally on the tongue tip, tongue dorsum, tongue back, upper lip, lower lip, and jaw. For the present analysis, only the data from the upper-lip and the tongue-tip sensors were analyzed. Reference sensors were placed on the bridge of the nose and the upper incisors. Their positions were used to correct for head movements and rotation of the data to the occlusal plane. The acoustic recordings were carried out with a DAT recorder and a Sennheiser MKH 20 P48 microphone. The distance between the microphone and the speaker's lips was about 30 cm. The acoustic signal was recorded with a sampling frequency of 48 kHz and was later downsampled to 24 kHz.

Sessions. In order to induce the speakers to change their articulatory strategy (and thus encourage the use of motor equivalent strategies), speakers were recorded in different conditions (see list below). For the recording made at the time of perturbation onset, speakers' auditory feedback was masked with white noise presented over headphones. This was done in the original study to investigate whether speakers could predict the acoustic outcome of a motor equivalence strategy—that is, whether they use a motor equivalence strategy with auditory feedback masked. The results were not clear because the horizontal articulator positions changed only slightly in this session as compared to the unperturbed session, as did the acoustic output. The following list summarizes all the sessions recorded in the experiment:

1. Day 1, Session 1: Unperturbed (without the prosthesis).
2. Day 1, Session 2: With the prosthesis in place and auditory feedback masked.
3. Day 1, Session 3: With the prosthesis in place and auditory feedback available.
4. Day 8: With the prosthesis in place after 1 week of adaptation.
5. Day 15, Session 1: With the prosthesis in place after 2 weeks of adaptation.
6. Day 15, Session 2: Unperturbed (without the prosthesis) after 2 weeks of adaptation.

In each session, each item was recorded 20 times.

Speech material. The target sounds /s/ and /ʃ/ were recorded in the nonsense words /ˈʃaxa/ and /ˈzasa/ spoken in a carrier phrase: *Ich sah Schacha an* ("I looked at /ˈʃaxa/."). Both sounds were recorded as part of the same corpus mixed with other material (other CVCV nonsense words consisting of all lingual sounds of German) in randomized order. The consonant /s/ was placed in noninitial position because it does not occur word-initially in German.

Calculation of acoustic distance. The acoustic contrast distance between /ʃ/ and /s/ was measured for the first unperturbed session only (20 repetitions per speaker). The consonants /ʃ/ and /s/ were segmented acoustically (friction onset to friction offset) in each utterance of the unperturbed session. The spectral center of gravity (COG) was calculated using the method of Forrest, Weismer, Milencovic, and Dougall (1988) with a pre-emphasis factor of 1, evaluated over an interval of 30 ms around the temporal midpoint of the fricative using a moving 6-ms Hamming window with 1 ms of overlap. The acoustic contrast distance between /s/ and /ʃ/ was calculated as the difference between mean COG of /s/ and /ʃ/.

Calculation of Mahalanobis distance between /s/ and /ʃ/. The Mahalanobis distance was measured for the unperturbed session only (20 repetitions per speaker). The purpose of this measure was to compare speakers' interphonemic variability with their intraphonemic variability. If, as suggested in earlier studies, high-acuity speakers really do produce clearer phonemic contrasts, their intraphonemic variability should be low as compared with their interphonemic variability.

In other words, for a high-acuity speaker, the distances between the centers of the goal regions should be relatively large and the sizes of the goal regions relatively small. For a low-acuity speaker, the distances between the centers of the goal regions should be relatively small and the sizes of the goal regions relatively large. The combination of these two factors would result in a more distinct acoustic difference between productions of the two sounds (greater clarity) for the high-acuity than for the low-acuity speaker.

The *Mahalanobis distance* is a measure of the distance of a single production (for example, an /s/ production) to the center of a sample of such productions (e.g., all /ʃ/ productions) expressed in units of variability within the sample (e.g., all /ʃ/ productions). Thus, the Mahalanobis distance is greater if the single production is far away from the center of the sample of productions but also if there is only little variability within the sample of productions.

A preliminary inspection of the spectra showed that inter- and intraphonemic variability existed not only with respect to the frequency in Hz of the main spectral peak (measured by the COG) but also with respect to skewness and kurtosis of the spectrum. Therefore, these higher spectral moments were calculated in the same way as the COG (cf. "Calculation of Acoustic Distance" section).

COG, skewness, and kurtosis values were z-normalized for each speaker in order to make interphonemic and intraphonemic variability comparable across speakers. Then the speaker-normalized Mahalanobis distances were computed in 3-D space (defined by COG, skewness, and kurtosis): Single /s/ productions were scaled in units of variability of the sample of /ʃ/ productions and, likewise, single /ʃ/ productions were scaled in units of variability of the sample of /s/ productions. Finally, for each speaker's sibilant contrast, a mean over all 40 Mahalanobis distances was calculated (from 20 /s/-tokens and 20 /ʃ/-tokens).

Word duration. As discussed in the introduction, we assumed a dependency of the use of motor equivalence on speech rate. During the experiment, it was noticed that two speakers spoke much more slowly than the remaining five, perhaps because of the unusual experimental design. Therefore, word duration was measured in order to assess speech rate. As discussed, the productions analyzed for calculating motor equivalence were part of the nonsense word /ˈʃaxa/, so the duration of this word was calculated as the time difference between the F2 offset of the final vowel /a/ and frication onset of the initial fricative.

Measurement of motor equivalence (R_{motequ}). If there is motor equivalence for a speaker in production of /ʃ/, she or he should covary horizontal tongue-tip and upper-lip position. Therefore, the horizontal positions of the tongue-tip and upper-lip sensors were measured over all sessions. There were 120 repetitions per speaker (six sessions with 20 repetitions in each session). For some speakers, it was noticed that the tongue-tip sensor was attached to a different location of the tongue during one session or another. For Speaker CM1, these were the sessions from Day 8 and Day 15; for Speaker AM2 and CF2, these were the sessions from Day 8. These sessions were removed from the sample. Furthermore, because of a broken sensor, there were no upper-lip data for speaker CF1 from Day 1. Also, five outliers (below or beyond the mean ±2 *SD*s) were removed from the data.

Pearson correlations between horizontal tongue-tip and upper-lip position were calculated, with each subject's data pooled across sessions. The resulting correlation coefficient was called R_{motequ}. This value designates the strength of the motor equivalence pattern. If a speaker has a clear covariation of tongue and lip position (a fronted tongue when the lip is fronted and a retracted tongue when the lip is retracted), this coefficient will be high. On the other hand, if a speaker has a coefficient close to 0, there is no covariation between tongue and lip and, thus, no motor equivalence.

Rate-adapted motor equivalence (R_{motequ_rate}). In order to account for the assumption that motor equivalence is influenced by speaking rate, the motor equivalence coefficient R_{motequ} was normalized by the speaking rate. In order to do so, R_{motequ} was multiplied by the mean word duration of the unperturbed session. The resulting value is called R_{motequ_rate}. For each subject, R_{motequ_rate} was higher for slow speech than for fast speech. The effect of speech rate on R_{motequ} was thereby minimized: The same R_{motequ} will result in a lower R_{motequ_rate} for a fast speaker (who presumably needs to use motor equivalence to a greater extent) than for a slow speaker (who does not need to use as much motor equivalence). For example, for speaker CF3, the fastest speaker, the mean word duration was 0.40 and R_{motequ} was 0.74. So, $R_{motequ\_rate} = 0.296$. Speaker AF1, the slowest speaker, had a mean word duration of 0.63 and a very low R_{motequ} of 0.27, resulting in $R_{motequ\_rate} = 0.17$.

Experiment 2: Auditory Acuity

The second experiment assessed the speakers' auditory acuity using a custom-synthesized /ˈasə/–/ˈaʃə/ continuum. In the first part of this experiment, speakers carried out a labeling test in order to determine their /s/–/ʃ/ boundary. Then, they performed a four-interval, two-alternative, forced-choice discrimination test around their phonemic boundary in order to assess their auditory acuity (Ghosh et al., 2010).

Synthesis of stimuli. The sibilant continuum was taken from the *said–shed* continuum used by Ghosh et al. (2010). This continuum was produced with the Klatt synthesizer (Klatt & Klatt, 1990) and comprised a gender-neutral sibilant continuum with 1,513 steps. The synthesis parameters were derived from natural utterances of the fricatives spoken by a male English speaker. Friction onset and friction offset were measured for each utterance. Within the frication interval, the amplitude at each prominent spectral peak was extracted. Then, a continuum was created by morphing between the sibilant segment parameters from the two endpoints (/s/ and /ʃ/). The fundamental frequency of every utterance was scaled to have a mean value of 165 Hz. This belongs to a region that corresponds to neither a prototypical male nor a prototypical female voice.

As a basis for the synthesis of the vowels, two German speakers, a man and a woman, recorded multiple repetitions of the words *Asse* and *Asche* in isolation. The fundamental frequency, the first five formants, and the bandwidths of the vowels were measured. Means were calculated across repetitions and across speakers. These values were then used as parameters in the Klatt synthesizer. The transitions from the vowel to the sibilant and the sibilant to the vowel were adjusted to assure natural-sounding continuity. Informal perceptual tests suggested that the syntheses sounded reasonably natural.

Procedure. Each speaker's category boundary between /s/ and /ʃ/ was determined by asking the subject to label a number of tokens as *Asse* or *Asche*. The tokens for labeling were selected from 11 equally spaced intervals between the endpoints of the /s–ʃ/ continuum. There were 10 trials for each token. For most of the subjects, the category boundary was established at the frequency values of the token that was not consistently labeled as either *Asse* or *Asche,* but where labeling occurred at chance level. If there was no such token, the boundary was established halfway between the two tokens where labeling swapped (i.e., between the two neighboring tokens of which one was consistently labeled *Asse* and the other one was consistently labeled *Asche*).

The subjects then performed a spectral discrimination test around their phonemic boundary as determined by the preceding procedure. This test used a four-interval forced-choice task in which the subject heard the sequence ABAA or AABA and indicated whether the second or third item was different from the rest. We used a four-interval test rather than a two-interval test because it has been shown that this procedure tests speakers' auditory processing (their auditory discrimination) rather than their phoneme labeling abilities (Gerrits & Schouten, 2004).

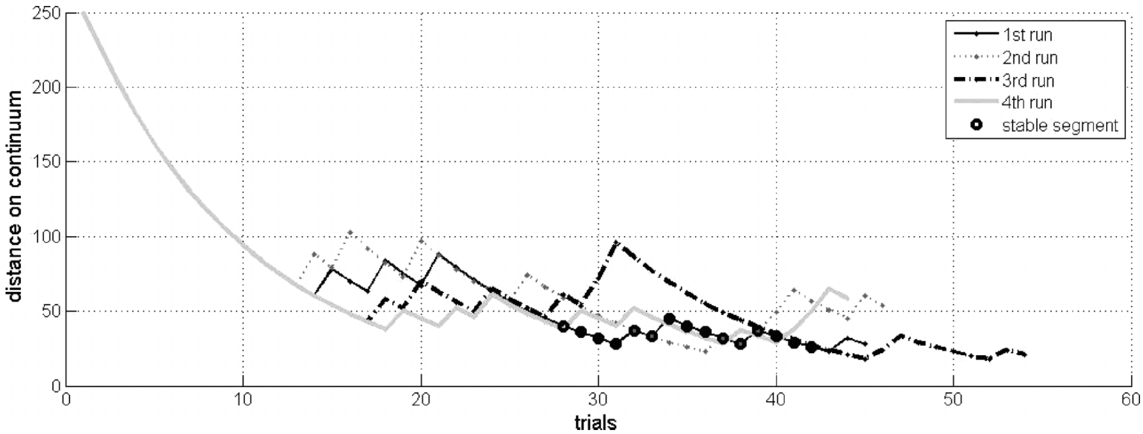
The interval was decreased by one step (10% of the separation between the stimuli used in the preceding trial) following a correct response and increased by three steps following an incorrect response. The test was terminated after 14 reversals or 80 trials, based on the average human sustained attention span, which is about 20 min (Kahneman, 1973).

There were four runs of this discrimination test for each speaker. Figure 2 shows an example for Subject CF3. The abscissa gives the trial number; the ordinate gives the auditory distance (in terms of numbers of stimuli along the continuum). Each of the lines in different styles and shades of grey shows the results of one run. In the beginning, it is evident that the distance between stimuli is reduced further and further with each trial following a correct response. Then, at about Trial 13, the curves start oscillating. In some runs, the oscillations are larger than in other runs, and even within a run there are passages that are more stable than others, possibly due to variation in the speaker's attention to the task. For the estimation of the speaker's JND parameter, the most stable sequence of trials was chosen (i.e., a sequence with only small oscillations around a mean). A first inspection of the data showed that listeners seemed to have rather short spans (around 15 trials) within which their behavior was relatively constant. In order to find the most stable sequence, squared deviations from the mean within each set of subsequent 15 trial intervals were calculated (sum of squared distances from the mean). A mean value was calculated over the 15 trials with the smallest deviation (cf. Figure 2, filled circles from Trial 28 to Trial 42). This value (i.e., the distance between two stimuli in units of the continuum) was converted into an acoustic measure (i.e., the distance in COG between the two stimuli). This parameter was called *JND.*

FIG2

Statistical analysis. The statistical analysis was carried out using R Version 2.9.0 (R Foundation for Statistical Computing, 2009). Because normality is hard to

Figure 2. Perception results for Subject CF3. Abscissa gives trials; ordinate gives difference between stimuli within a trial in continuum units. Different line colors and styles show different runs of the experiment (see key). Black circles show the 15-trial interval with the smallest oscillations around a mean. The mean over this interval was used for estimating the JND.



assess for a small sample such as the one used in this study, nonparametric tests were used, namely, permutation tests calculating correlations with 1,000 random pairings (Edgington, 1995). The α-level was set to .05 (one-tailed). The tests were calculated for the following relations: acoustic distance and JND, mean Mahalanobis distance and JND, R_{motequ} and JND, and R_{motequ_rate} and JND.

Results

JND and Acoustic Distance

FIG3 Figure 3a shows the relation between auditory acuity (JND) and the distance between the average spectral means of /s/ and /ʃ/. The abscissa shows the JND and the ordinate gives the mean acoustic distance between /s/ and /ʃ/ for the respective speaker. According to our hypothesis, there should be a negative correlation. Speakers with a low JND (high-acuity speakers) should have a greater acoustic distance than low-acuity speakers. TBL1 Table 1 gives the results of the permutation tests. The correlation coefficient for this relation is –0.759 and the significance is .037. This result is consistent with the hypothesis that the JND influences phonemic distance; the lower the JND, the greater the contrast distance.

Table 1. Results of the permutation test.

Variable	JND
Acoustic distance	$r = -.759$, $p = .037$*
Mean Mahalanobis distance	$r = -.845$, $p = .037$*
R_{motequ}	$r = -.664$, $p = .056$, *ns*
R_{motequ_rate}	$r = -.770$, $p = .038$*

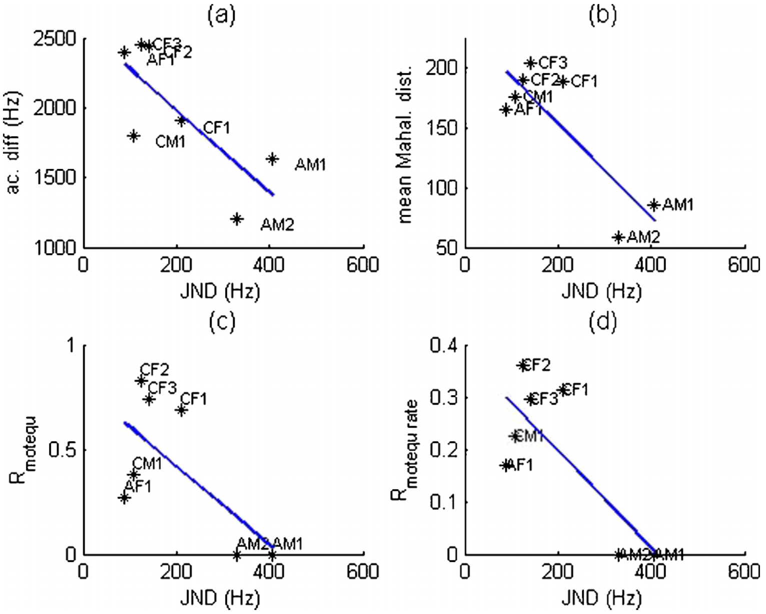
Note. The first value reflects the correlation coefficient and the second value reflects the significance. JND = just noticeable difference.

*$p < .05$.

Mahalanobis Distance Between /s/ and /ʃ/ and JND

Figure 3b shows the relation between JND and the mean Mahalanobis distance between the two fricatives. The fact that all Mahalanobis distances are greater than 1 shows that there was no overlap in 3-D space between the two categories. According to our hypothesis the distance should be greater for high-acuity speakers (low JND) than for low-acuity speakers (high JND). This relation can be seen in the figure. The correlation is significant (cf. Table 1).

Figure 3. (a) Relation between JND and the /s–ʃ/ distance. ac. diff. = acoustic difference. Abscissa gives JND (in Hz); ordinate gives the acoustic distance between /s/ and /ʃ/ (in Hz). Low JND = high-acuity speakers; high JND = low-acuity speakers. (b) Relation between JND (abscissa) and mean Mahalanobis distance between /s/ and /ʃ/ (ordinate). (c) Relation between JND (abscissa) and R_{motequ} (ordinate). Low R_{motequ} = no motor equivalence; high R_{motequ} = clear motor equivalence. (d) Relation between JND (abscissa) and R_{motequ_rate} (ordinate).



If one compares the Mahalanobis distance (Figure 3b) with the mean acoustic distance (Figure 3a), one can see that there are some differences. The mean Mahalanobis distances are bimodally distributed: There are two speakers with a very low Mahalanobis distance and other speakers with a very high Mahalanobis distance. In contrast to that, the acoustic difference of the COG is more equally distributed. A reason for this could be that speakers create the difference by individual variation among the acoustic parameters (i.e., COG, skewness, kurtosis).

Word Duration

FIG4 Figure 4 shows the results for the duration of the nonsense word /ˈʃaxa/ for the different speakers. The abscissa gives the speaker; the ordinate gives the word duration in seconds. The word duration is about the same for speakers AM1, AM2, CF1, CF2, and CF3, but much longer for speakers AF1 and CM1.

R_{motequ}

TBL2 Table 2 shows the correlation coefficients for the relation horizontal upper-lip position–horizontal tongue-tip position. It is evident that five speakers (AF1, CM1, CF1, CF2, CF3) had significant positive correlations; one speaker (AM1) had a significant negative correlation and one speaker (AM2) had no significant correlation.

A negative correlation between tongue position and lip position means that there is more lip protrusion when the tongue is retracted. Since it would not be reasonable to expect a speaker with a strong negative correlation to have even worse auditory acuity than a speaker with only a weak negative correlation, the two speakers with a negative coefficient were assigned an R_{motequ} of 0 (i.e., in subsequent analysis, the R_{motequ} of speakers AM1 and AM2 is taken as 0, cf. Figures 3c and 3d).

Figure 4. Mean word duration in seconds for different speakers. Boxes show lower quartile, median, upper quartile; whiskers end at 1.5 quartiles. Boxes and whiskers cover about 99% of the data.

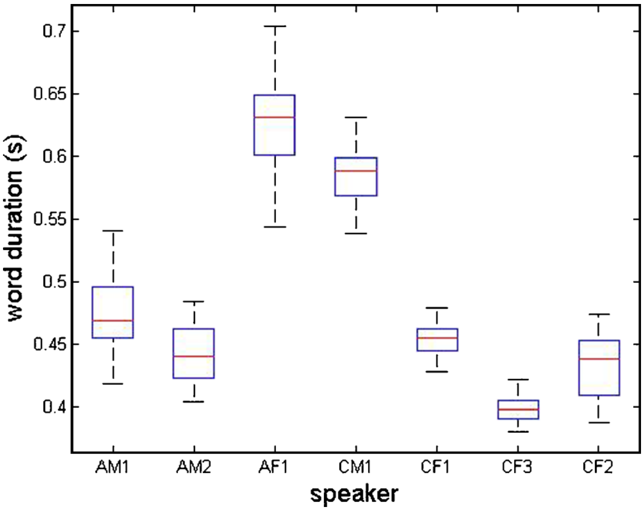


Table 2. Correlation between horizontal tongue-tip position and horizontal upper-lip position (R_{motequ}).

Speaker	R_{motequ} (*p*)[a]	N[b]
AM1	–0.40 (.000)	119
AM2	–0.08 (.439)	100
AF1	0.27 (.003)	120
CM1	0.38 (.002)	60
CF1	0.69 (.000)	58
CF2	0.83 (.000)	100
CF3	0.74 (.000)	118

[a]Correlation coefficients and *p* values for the relation between tongue position and lip position. [b]Number of repetitions available for the analysis. Variation in *N* among speakers is due to removal of sessions with misplaced sensors.

R_{motequ}/R_{motequ_rate} and JND

Figure 3c shows the relation between R_{motequ} (the motor equivalence coefficient) and JND. Following our hypothesis there should be a negative correlation between the two parameters: Speakers with a high R_{motequ} (clear motor equivalence pattern) should have a low JND (high acuity); speakers with a low R_{motequ} (no motor equivalence) should have a high JND (low acuity). There is a tendency for speakers with a high JND to have a low R_{motequ} and for speakers with a low JND to have a high R_{motequ}. However, the correlation is not significant in the permutation test (cf. Table 1).

Looking at individual speakers one can see that there are two speakers with very low JNDs who, however, do not have a clear motor-equivalence pattern (AF1 and CM1). These two speakers also had very long word durations (cf. Figure 4).

As stated in the introduction, one can expect a trade-off between motor equivalence and speech rate: Because motor equivalence is a means of reducing articulatory effort, and because reduction of articulatory effort is more vital in fast than in slow speech, there should be more motor equivalence in fast than in slow speech. This could explain the behavior of speakers AF1 and CM1: Because of their slow speech, they don't need to use much motor equivalence. Given their high auditory acuity, they probably would show more motor equivalence if they spoke faster.

R_{motequ_rate} accounts for the differences in duration and their effect on motor equivalence. The permutation test between R_{motequ_rate} and JND gives a significant result ($r = -.770$, $p = .038$). Thus, auditory acuity has an influence on the use of motor equivalence in adaptation if word duration is taken into account.

Discussion

During the production of the fricative /ʃ/, speakers can use a motor-equivalent trading relation in varying the position of the tongue constriction and the amount of lip protrusion to maintain a relatively stable front-cavity length and resulting acoustic output. The present study investigated the relationship among the degree of motor equivalence during /ʃ/, auditory acuity for the sibilant contrast, and produced acoustic contrast distance in seven speakers. Two hypotheses were established. The first was that high-acuity speakers should use motor equivalence to a greater degree than low-acuity speakers in reaching their smaller perceptual goal regions. The second hypothesis was that high-acuity speakers should produce clearer phonemic contrasts between /s/ and /ʃ/.

Motor equivalence in /ʃ/ was measured from seven speakers as they adapted to an artificial palate worn over a period of two weeks. The speakers' horizontal-lip and tongue-tip positions during /ʃ/ were measured and correlations between these two articulatory variables were calculated, resulting in a measure of motor equivalence, R_{motequ}. Following earlier findings on speech motor control (Nelson, 1983; Perkell et al., 2002) it was expected that speech rate would have an influence on motor equivalence. If motor equivalence is regarded as a mechanism to reduce articulatory effort, it should play less of a role in slow speech (where the articulatory effort is already low) than in fast speech. In order to account for this inference, the motor-equivalence coefficient R_{motequ} was multiplied with the speaker's word duration (R_{motequ_rate}).

Phonemic contrast distance was measured in two ways, first, as the difference in Hz between the mean spectral COG of unperturbed /s/ and /ʃ/ productions, and second, in order to account for differences in inter- and intraphonemic variability within a sound category, as the mean Mahalanobis distance between /ʃ/ and /s/ in a 3-D space defined by the z-normalized spectral center of gravity, skewness, and kurtosis.

Auditory acuity was assessed in the same speakers with perceptual discrimination test on an /s–ʃ/ continuum.

The results are consistent with both hypotheses. High-acuity speakers showed greater phonemic distances between /s/ and /ʃ/ (i.e., clearer contrasts). Both the correlation between JND and acoustic distance and the one between JND and Mahalanobis distance were significant. Taking into account differences in word duration, there was a significant correlation across subjects between the measure of motor equivalence and acuity. High-acuity speakers used motor equivalence to a greater extent than low-acuity speakers. Some speakers with high auditory acuity spoke slower than others, and these speakers used less motor equivalence, presumably because motor equivalence is a means of conserving articulatory effort and would therefore be more evident in faster than in slower speech.

In the 3-D space defined by spectral COG, skewness, and kurtosis, the regions of the two fricatives did not overlap for any of the speakers. Therefore, listeners should not have difficulty classifying a speaker's production as one or the other fricative even for the low-acuity speakers whose contrast distances between the fricatives were smaller. These observations then lead to a question: Why should high-acuity speakers then produce these greater distances? The high-acuity speakers' greater contrast distance cannot be explained simply by the listeners' needs. A better explanation can be found in the development of the speakers' perceptual regions during speech acquisition. A child acquiring speech will learn that it is advantageous to be as intelligible as possible. Thus, the child will try to produce speech sounds with the greatest possible contrast. As a high-acuity speaker the child will be able to perceive smaller acoustic differences between productions and reject some productions as produced badly whereas such sounds would be acceptable for a low-acuity speaker (Perkell, in press).

As stated in the introduction, motor equivalence generally is difficult to observe. The correlation coefficients given in Perkell et al. (1993) for lip rounding and tongue-body raising in many repetitions of the vowel /u/ are between –0.21 and –0.47, which is somewhat smaller than the numbers found in our data. A possible explanation for this difference could be that motor equivalence plays a more important role under perturbation, where the speaker has to find a new way to produce the sound in an efficient manner. The speakers might try a variety of articulations, which could result in the correlations found in the study.

Savariaux et al. (1995) found large articulatory changes in /u/ when they perturbed the lip movement with a lip tube. Some speakers retracted the tongue towards the velo-pharyngeal region instead of producing a velar constriction that they normally used to produce an unperturbed /u/. This result is in line with the assumption that more motor equivalence can be found when speech is perturbed than when it is not.

Westbury et al. (1998) found motor equivalence across different contexts, whereas within a context speakers tended to maintain a particular way of producing /r/. This observation is consistent with our assumption that without perturbation speakers have a preferred, habitual way to produce the sound, one they probably have selected because it involves a minimum of effort.

Additional research should be carried out with the aim of understanding the determinants of motor equivalence in speech production. The relationship between speech rate and motor equivalence, for example, is still not well understood. Our results lead to the inference that speakers' use of motor-equivalence strategies depends on their ability to hear small acoustic differences; only if they are able to perceive differences in the acoustic output will they try to correct for those differences by using motor equivalence. The results can therefore help to explain the interspeaker variability found in a number of studies.

Acknowledgments

This research was supported by a DAAD (German Academic Exchange Service) grant to the first author for carrying out postdoctoral research and by National Institute on Deafness and Other Communication Disorders Grant R01DC00925 (J. Perkell, P.I.). Thanks to Jörg Dreyer and Melanie Weirich for carrying out perception tests.

References

Benguerel, A. P., & Cowan, H. (1974). Coarticulation of upper lip protrusion in French. *Phonetica, 30,* 41–55.

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

Brunner, J., & Hoole, P. (2009). *Motor equivalent strategies in the production of German /ʃ/ under perturbation*. Manuscript submitted for publication.

Delattre, P., & Freeman, D. (1968). A dialect study of American R's by x-ray motion picture. *Linguistics, 44,* 29–68.

Edgington, E. S. (1995). *Randomization tests*. New York, NY: Marcel Dekker.

Edwards, J. (1985). Contextual effects on lingual-mandibular coordination. *The Journal of the Acoustical Society of America, 78,* 1944–1948.

Fant, G. (1960). *Acoustic theory of speech production*. The Hague, the Netherlands: Mouton.

Folkins, J. W., & Zimmermann, G. N. (1982). Lip and jaw interaction during speech: Responses to perturbation of lower-lip movement prior to bilabial closure. *The Journal of the Acoustical Society of America, 71,* 1225–1233.

Forrest, K., Weismer, G., Milencovic, P., & Dougall, R. N. (1988). Statistical analysis of word-initial voiceless obstruents: Preliminary data. *The Journal of the Acoustical Society of America, 84,* 115–123.

Fowler, C. A., & Saltzman, E. (1993). Coordination and coarticulation in speech production. *Language and Speech, 36,* 171–195.

Gerrits, E., & Schouten, M. E. H. (2004). Categorical perception depends on the discrimination task. *Perception & Psychophysics, 66,* 363–376.

Ghosh, S. S., Matthies, M. L., Maas, E., Hanson, A., Tiede, M., Ménard, L., … Perkell, J. S. (2010). An investigation of the relation between sibilant production and somatosensory and auditory acuity. *The Journal of the Acoustical Society of America, 128,* 3079–3087.

Gracco, V. L., & Abbs, J. H. (1985). Dynamic control of the perioral system during speech: Kinematic analyses of autogenic and nonautogenic sensorimotor processes. *Journal of Neurophysiology, 54,* 418–432.

Guenther, F. H. (1994). A neural network model of speech acquisition and motor equivalent speech production. *Biological Cybernetics, 72,* 43–53.

Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review, 102,* 594–621.

Guenther, F. H., Espy-Wilson, C. Y., Boyce, S. E., Matthies, M. L., Zandipour, M., & Perkell, J. S. (1997). Articulatory tradeoffs reduce acoustic variability during American English /r/ production. *The Journal of the Acoustical Society of America, 105,* 2854–2865.

Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language, 96,* 280–301.

Guenther, F. H., Hampson, M., & Johnson, D. (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychological Review, 105,* 611–633.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kelso, J. A. S., & Tuller, B. (1983). "Compensatory articulation" under conditions of reduced afferent information: A dynamic formulation. *Journal of Speech and Hearing Research, 26,* 217–224.

Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America, 87,* 820–857.

Lindblom, B., Lubker, J., & Gay, T. (1979). Formant frequencies of some fixed-mandible vowels and a model of speech motor programming by predictive simulation. *Journal of Phonetics, 7,* 147–161.

McFarland, D. H., & Baum, S. R. (1995). Incomplete compensation to articulatory perturbation. *The Journal of the Acoustical Society of America, 97,* 1865–1873.

Nelson, W. L. (1983). Physical principles for economies of skilled movements. *Biological Cybernetics, 46,* 135–147.

Newman, R. S. (2003). Using links between speech perception and speech production to evaluate different acoustic metrics: A preliminary report. *The Journal of the Acoustical Society of America, 113,* 2850–2860.

Perkell, J. S. (in press). Movement goals and feedback and feedforward control mechanisms in speech production. *Journal of Neurolinguistics*.

Perkell, J., Guenther, F., Lane, H., Matthies, M., Perrier, P., Vick, J., … Zandipour, M. (2000). A theory of speech motor control and supporting data from speakers with normal hearing and with profound hearing loss. *Journal of Phonetics, 28,* 233–272.

Perkell, J. S., Guenther, F. H., Lane, H., Matthies, M. L., Stockmann, E., Tiede, M., & Zandipour, M. (2004). The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts. *The Journal of the Acoustical Society of America, 116,* 2338–44.

Perkell, J. S., Matthies, M. L., Svirsky, M. A., & Jordan, M. I. (1993). Trading relations between tongue-body raising and lip rounding in production of the vowel /u/: A pilot motor equivalence study. *The Journal of the Acoustical Society of America, 93,* 2948–2961.

Perkell, J. S., Matthies, M. L., Tiede, M., Lane, H., Zandipour, M., Marrone, N., . . . Guenther, F. H. (2004). The distinctness of speakers' /s/-/ʃ/ contrast is related to their auditory discrimination and use of an articulatory saturation effect. *Journal of Speech, Language, and Hearing Research, 47,* 1259–1269.

Perkell, J. S., Zandipour, M., Matthies, M. L., & Lane, H. (2002). Economy of effort in different speaking conditions. I. A preliminary study of intersubject differences and modeling issues. *The Journal of the Acoustical Society of America, 112,* 1627–1641.

R Foundation for Statistical Computing. (2009). R (Version 2.9.0) [Computer software]. Retrieved from http://www.r-project.org/.

Savariaux, C., Perrier, P., & Orliaguet, J.-P. (1995). Compensation strategies for the perturbation of the rounded vowel [u] using a lip-tube: A study of the control space in speech production. *The Journal of the Acoustical Society of America, 98,* 2428–2442.

Savariaux, C., Perrier, P., Orliaguet, J.-P., & Schwartz, J.-L. (1999). Compensation strategies for the perturbation of French [u] using a lip tube. II. Perceptual analysis. *The Journal of the Acoustical Society of America, 106,* 381–393.

Sussman, H. M., & Westbury, J. (1981). The effects of antagonistic gestures on temporal and amplitude parameters of anticipatory labial coarticulation. *Journal of Speech and Hearing Research, 24,* 16–24.

Westbury, J., Hashi, M., & Lindstrom, M. (1998). Differences among speakers in lingual articulation of American English /r/. *Speech Communication, 26,* 203–226.

Zhou, X., Espy-Wilson, C., Boyce, S., Tiede, M., Holland, C., & Choe, A. (2008). A magnetic resonance imaging-based articulatory and acoustic study of "retroflex" and "bunched" American English /r/. *The Journal of the Acoustical Society of America, 123,* 4466–4481.

AUTHOR QUERY

No query.