Familiarization Effects on Consonant Intelligibility in Dysarthric Speech

Heejin Kim

Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, Ill., USA

Key Words

Dysarthria · Intelligibility · Speech perception

Abstract

Background/Aims: This study investigates the effects of familiarization on naïve listeners' ability to identify consonants in dysarthric speech. **Methods:** A total of 120 listeners (30 listeners/speaker) participated in experiments over a 6-week period. Listeners were randomly assigned to one of the three familiarization conditions: a passive condition in which listeners heard audio recordings of words, an active condition in which listeners heard audio recordings of words while viewing the written material of words, and a control condition in which listeners had no exposure to the audio signal prior to identification tasks. **Results:** Familiarization improved naïve listeners' ability to identify consonants produced by a speaker with dysarthria. The active familiarization method exhibited an advantage over the other conditions, in terms of the magnitude and rapidness of improvement. One-month delayed test scores were higher than pre-familiarization scores, but the advantage of active familiarization was not present for all speakers. **Conclusion:** This study supports familiarization benefits in enhancing consonant intelligibility in dysarthria and suggests that perceptual learning mechanisms be harnessed for developing effective listener-oriented intervention techniques in the

management of dysarthria. Current findings call for further research on a familiarization protocol that can subserve segmental learning with maximum efficacy.

© 2016 S. Karger AG, Basel

Introduction

Familiarization, a method by which listeners receive a brief exposure to atypical speech, has demonstrated potential as an intelligibility enhancement technique for speakers with dysarthria. The majority of familiarization studies in dysarthria reported that naïve listeners improved their ability to understand the speech when they underwent a familiarization phase prior to intelligibility assessments compared to nonfamiliarized listeners [1–9]. Two studies [10, 11], however, reported no significant effect of familiarization, which appears to be due to the methods employed. Unlike all other studies, Yorkston and Beukelman [10] employed experienced listeners who might have already reached their ceiling point. Garcia and Cannito [11] used a passive familiarization method – that is, listeners were presented only with an audio signal of dysarthric speech – while other studies used an active familiarization condition in which listeners were familiarized both with an audio signal and a written transcript. According to Borrie et al. [2] and Kim and Nanney

[5], an active familiarization condition yielded more robust intelligibility gains than passive familiarization, suggesting that the passive familiarization condition used by Garcia and Cannito [11] was not sufficiently effective to improve intelligibility. Another type of familiarization method was tested in Hustad and Cahill [4], in which listeners underwent 4 repeated experimental tasks with no explicit familiarization. Listeners exhibited improvements in intelligibility, but the point at which the benefit of familiarization became significant was influenced by the severity of speakers.

Familiarization effects have been evidenced not only with dysarthria but also with a variety of atypical speech: e.g., nonnative-accented speech [12–14], speech produced by individuals with hearing impairment [15–17], and artificially created or altered speech such as synthesized voice, noise-vocoded speech, and time-compressed speech [18–25]. These recurring observations of familiarization effects in the field of speech perception underscore the flexibility of speech perception and point to a cognitive-perceptual process, termed perceptual learning, which listeners engage when they are faced with atypical speech. By this process, listeners are capable of recognizing speech that sounds deviant from what they know as normal and can recalibrate existing categories, with the result of an improved ability to understand atypical speech that was initially difficult to perceive [26, 27]. Aiming to evaluate the potential rehabilitative gain through familiarization in dysarthria, researchers have shown growing interest in exploiting perceptual learning mechanisms as a foundational principle for developing an optimal protocol of familiarization [3].

In order to understand the mechanisms that underlie familiarization-induced perceptual learning, it is essential to investigate the learning sources and the factors that facilitate or impede the learning. As speech perception is a complex process involving interactions among numerous paralinguistic and linguistic features, attempts have been made to factor out the relative benefits of various features in familiarization-induced perceptual change. There is evidence that a paralinguistic feature such as a speaker's voice plays a key role in perceptual processing of speech in noise [28, 29], noise-vocoded speech [30], and dysarthric speech [31]. In addition, the benefit of prosodic information in perceptual learning is supported by studies of time-compressed speech [24, 25], while a conflicting finding has been reported [8, 9] in which the presence of sentence-level prosody in paragraphs did not result in higher intelligibility scores compared to the isolated word condition.

Segmental benefits in perceptual learning have also been discussed in many studies. A postulation is shared in the field that the source of perceptual benefit in improved word intelligibility is primarily at the segmental level of perceptual reorganization by which listeners adjust the mapping between acoustic-phonetic information and the phonemic representation in the language [3, 18, 32–34]. Changes in listeners' segmental perception after familiarization training have been reported for vowels in nonnative speech [35, 36], and for consonants in synthesized speech [18] and in normal, native speech [37–39]. Indirect support can be found concerning dysarthric speech in Liss et al. [6], who examined lexical boundary errors for familiarized versus nonfamiliarized listeners. A major finding was that two listener groups did not differ in terms of the lexical boundary error types, although familiarized listeners exhibited improved word intelligibility compared to nonfamiliarized listeners. The investigators concluded that the intelligibility improvement could not be attributed predominantly to prosodic knowledge of strong/weak syllables, but rather to improved performance at the segmental level [6]. Borrie et al. [2] and Spitzer et al. [7] extended support for the segmental benefits of familiarization based on the finding that the familiarized listeners showed more word substitution errors that bore phonemic resemblance to the target compared to nonfamiliarized listeners.

While the above studies in dysarthria indicated the possible segmental benefits of familiarization, it should be noted that speech perception is an interactive process involving the bidirectional contribution of bottom-up (low-level acoustic) and top-down (high-level lexical or semantic) flow of information to phoneme categorization [3, 27]. In particular, lexical knowledge aids phoneme identification, especially when the signal is degraded [27, 40]. Therefore, interpretation of results drawn from word transcription tasks warrants caution, unless the transcription tasks used non-words [as in 18, 35–39]. As Francis et al. [18] stated, investigation is needed to determine whether perceptual adjustments are made at a lexical level or at a segmental level, to discover how listeners can best be trained to understand atypical speech.

The purpose of this current study is to investigate the effects of familiarization on consonant intelligibility in dysarthria using naïve listeners. Unlike prior studies that assessed segmental benefits of familiarization in dysarthria based on word transcription tasks that require lexical decision, this study conducted a direct measure of consonantal intelligibility based on isolated consonant identification tasks. This is our second study in a series

that explores familiarization effects on dysarthric speech intelligibility secondary to cerebral palsy. In the first work [5], active familiarization exhibited intelligibility improvement at the word level. This current study examines whether a familiarization effect is manifest also at the segmental level. As in our previous study [5], speaker-specific familiarization was investigated through 4 familiarization phases as well as a 1-month delayed test, comparing three experimental groups of naïve listeners: a passive-familiarization group (audio only, no written transcript), an active-familiarization group (audio plus written transcript), and a no-explicit-familiarization group. The no-explicit-familiarization group was similar to the experimental group in Hustad and Cahill's study [4], in that this group did not receive any separate familiarization training and only underwent 4 sequential tasks. As discussed in Hustad and Cahill [4], it was not clear how much of the observed effect was actually due to familiarization with the speech rather than familiarization with the experimental tasks because their experimental design did not include comparison groups. Inclusion of the three groups in the current investigation allows examination of whether an additional effect in passive and/or active groups exists, compared to the no-explicit-familiarization group. If the degree of consonant intelligibility enhancement turns out to be greater in the passive and/or active groups compared to the no-explicit-familiarization group, the extra gain is due to the familiarization condition that was available only for these groups. Similarly, if we find extra gains in the active groups compared to the passive group, it would be clearly due to the presence of written materials as a supplement to the audio signal.

Findings in this study will contribute to improved understanding of perceptual learning sources. In addition, findings will be instructive in terms of the clinical utilization of familiarization for dysarthria management. The following research questions were addressed: (1) does familiarization improve consonant intelligibility in dysarthric speech?, (2) does the gain size in consonant intelligibility vary depending on familiarization condition?, and (3) are familiarization effects on consonant intelligibility retained long-term?

Materials and Methods

Speakers

Speakers were 4 American-English native speakers diagnosed with cerebral palsy. Among the 4 speakers, 3 speakers were the same as in our first work on word intelligibility [5]. One

female speaker was replaced with a new female speaker because her production errors were mostly in vowels. All 4 speakers were a part of the Universal Access (UA) corpus, which was developed as a resource for developing automatic speech recognition systems for dysarthric speech [41]. The overall intelligibility levels of the speakers were estimated based on naïve listeners' single word transcription tasks: 5 listeners were assigned to each speaker, each listener transcribed 225 words, and the percentage correct was then averaged across the 5 listeners to yield each speaker's intelligibility. All 4 speakers had intelligibility scores in the moderate range, to avoid possible ceiling and floor effects. Their age, gender, and intelligibility were as follows: Speaker 1 (58; male; 28%), Speaker 2 (18; male; 39%), Speaker 3 (21; male; 59%), and Speaker 4 (18; female; 62%). All 4 speakers were from northern or central Illinois. Three speakers (Speakers 1, 2, and 3) had spastic dysarthria, and 1 speaker, Speaker 4, had athetoid dysarthria.

Listeners

Listeners were recruited using the following criteria: (a) native speakers of American English, (b) between 18 and 40 years old, (c) no identified language, learning, or cognitive disability per self-report, (d) no more than incidental experience with persons with speech disorders, (e) grew up in the North Midland dialect region, and (f) passed a hearing screening at 20 dB for 250, 500, 1,000, 2,000, 4,000, 6,000, and 8,000 Hz. Thirty listeners were assigned to each speaker and were divided randomly into 3 groups of 10. Group A was familiarized using only the audio signal (passive familiarization), Group B was familiarized using both the audio signal and a written transcript (active familiarization), and Group C received no explicit familiarization (controls).

Stimuli

The stimuli were constructed from the UA database [41]. Two sets of stimuli were constructed: one for familiarization phases and one for consonant identification tasks. The stimuli for the familiarization phase were real words in American English and consisted of a speaker's 2 different productions of 32 distinct words, yielding a total of 64 tokens per speaker. The 16 consonants of interest, /p, b, t, d, k, g, m, n, s, z, ʃ, h, l, r, w, j/, all occurred in the word initial position. Each consonant occurred in 4 words (2 different productions of 2 distinct words), permitting a balanced number of consonants (16 consonants $\times$ 4 = 64 words). All words were mono- or bisyllabic and had lexical stress on the first syllable, controlling possible confounding effects of linguistic complexity and lexical stress. Peak amplitude of each stimulus was normalized to a level of 69 dB SPL.

For consonant identification tasks, consonant stimuli were excised from whole words. The whole words that consonants were spliced from included the same 32 words as the familiarization stimuli, but from different productions by the same speaker. In addition, 31 novel words that were not used in familiarization provided consonant stimuli, resulting in a total of 63 identification stimuli: 63, instead of 64, were constructed since there were only 3 /z/ tokens in the database that satisfied the criteria listed above. Each consonant occurred 4 times, except /z/, which occurred 3 times. Consonant stimuli included the consonant interval and 50 ms of the following vowel, and they were extracted using Praat [42].

Procedure

Experiments took place in a sound-treated room. Listeners were seated directly in front of a laptop computer in which a custom-built experimental interface was run to play stimuli and collect listeners' transcriptions. Audio recordings of stimuli were presented via a loudspeaker approximately 2 feet away from the listeners. The peak output level of stimulus material was approximately 67 dB from where listeners were seated. Listeners participated in 3 experimental sessions over a 6-week period. Before starting each day of the experiment, listeners completed a practice round to get acquainted with the experimental interface and procedures. Listeners were informed that they would hear the recordings of speech produced by a single individual with a speech disorder. They were also told that all words would be real words in American English and that all consonants were excised from the initial position of the real words. Listeners completed both word transcription and consonant identification tasks on the same day, but since the focus of this study is on consonant intelligibility, described below are mainly the procedures of familiarization and consonant identification tasks.

During Session 1 (pre-familiarization test), all listeners completed one round of consonant identification tasks. They listened to one of the consonants in American English (annotated on the screen as *p, b, t, d, k, g, m, n, s, z, sh, h, l, r, w, y*) and were asked to click the appropriate consonant button matching what they thought the speaker produced. Stimuli were presented one at a time and only once. One week later, listeners returned for Session 2 to complete 4 rounds of identification tasks. Before they performed each round of identification task, listeners in Group A (passive familiarization) had a familiarization phase in which they heard a series of 64 words produced by their assigned speaker. Listeners in Group B (active familiarization) heard words while reading a corresponding orthographic representation of the target word on the computer screen. Listeners in Group C (controls) received no familiarization phase. Approximately one month after Session 2, listeners returned to complete Session 3 (long-term, delayed test) that consisted of the same tasks as in Session 1.

Analysis

Identified consonants were scored as either correct or incorrect, and the percentage of correct responses per listener was calculated using a customized computer program. Two-way mixed ANOVA analysis was conducted separately for each speaker. The dependent measure was 'Identification Accuracy' (i.e., the percentage of correct consonants). The between-group independent factor was 'Group', which represented different familiarization conditions (i.e., Group A; Group B; Group C). The repeated measure was 'Session' and its categories were Session 1, Session 2 – first round, Session 2 – second round, Session 2 – third round, Session 2 – fourth round, and Session 3 (henceforth referred to as Session 1, Session 2-1, Session 2-2, Session 2-3, Session 2-4, and Session 3). The magnitude of difference between means was assessed using an effect size measure, η^2 (generalized eta squared), which provides comparability across between- and within-subject designs [43]. By this method, an effect size of 0.26 is considered large, 0.13 is considered medium, and 0.02 is considered small. Subsequent post hoc analyses were conducted, using a Tukey adjustment method that controls type I error rate and adjusts *p* values accordingly.

Prior to main analyses, a one-way ANOVA was conducted on Session 1 scores to assure that listeners' base performance before familiarization was equivalent across different experimental groups for a given speaker. The ANOVA results confirmed the uniformity of the experimental groups (Speaker 1, $p = 0.272$; Speaker 2, $p = 0.172$; Speaker 3, $p = 0.205$; Speaker 4, $p = 0.265$).

Results

Figure 1 illustrates the mean scores of each session for the three experimental groups. Each panel represents a speaker. For all speakers, all three experimental groups showed higher Session 2 scores than Session 1 scores. The intelligibility scores decreased in Session 3 compared to Session 2, but they were still higher than Session 1 scores for all experimental groups across all speakers. In addition, listeners in Group B (active familiarization) exhibited higher scores compared to Group A (passive familiarization) and Group C (controls) in all of the Session 2 rounds and Session 3 for all speakers.

Two-way mixed ANOVA analyses indicated a significant main effect of Session for all speakers [Speaker 1, $F(5, 130) = 16.71$, $\eta^2 = 0.195$, $p < 0.0001$; Speaker 2, $F(5, 135) = 23.27$, $\eta^2 = 0.331$, $p < 0.0001$; Speaker 3, $F(5, 135) = 10.17$, $\eta^2 = 0.130$, $p < 0.0001$; Speaker 4, $F(5, 135) = 42.92$, $\eta^2 = 0.382$, $p < 0.0001$]. In addition, the interaction Group $\times$ Session was significant except for Speaker 3 ($p = 0.499$), indicating that Session effect varied depending on Group [Speaker 1, $F(10, 130) = 2.81$, $\eta^2 = 0.075$, $p = 0.003$; Speaker 2, $F(10, 135) = 2.42$, $\eta^2 = 0.093$, $p = 0.011$; Speaker 4, $F(10, 135) = 3.77$, $\eta^2 = 0.098$, $p < 0.0001$]. The main effect of Group was not significant for any speakers (Speaker 1, $p = 0.188$; Speaker 2, $p = 0.151$; Speaker 3, $p = 0.234$; Speaker 4, $p = 0.072$).

Follow-up analyses on the interaction using Tukey post hoc analysis revealed that for Speaker 1, Group A's Session 1 score was lower than Session 2-3 ($p = 0.0447$), while Group B's Session 1 score was lower than all of its following sessions: Session 2-1 ($p = 0.0372$), Session 2-2 ($p = 0.0065$), Session 2-3 ($p < 0.001$), Session 2-4 ($p < 0.001$), and Session 3 ($p = 0.0134$). For Group C, no comparison was significant. Similarly, for Speaker 2, Group A's Session 1 was lower than Session 2-3 ($p = 0.0134$) and Session 2-4 ($p = 0.0042$), while Group B's Session was lower than all of the subsequent sessions: Session 2-1 ($p < 0.001$), Session 2-2 ($p = 0.0048$), Session 2-3 ($p < 0.001$), Session 2-4 ($p < 0.001$), and Session 3 ($p < 0.001$). For Group C, the Session 1 score was lower than Session 2-3 ($p = 0.0074$), Session 2-4 ($p = 0.0467$), and Session 3 ($p = 0.0074$). For Speaker 4, Group A's Session 1 was low-

Fig. 1. Mean identification accuracy scores plotted against each session for the three experimental groups: * = Group A (passive familiarization); O = Group B (active familiarization); Δ = Group C (control).

er than Session 2-2 ($p = 0.0302$), Session 2-3 ($p = 0.0080$) and Session 2-4 ($p < 0.001$), while Group B's Session 1 was lower than all the following sessions: Session 2-1 ($p = 0.0018$), Session 2-2 ($p < 0.001$), Session 2-3 ($p < 0.001$), Session 2-4 ($p < 0.001$), and Session 3 ($p < 0.001$). For Group C, the Session 1 score was lower than Session 2-3 ($p = 0.0131$), Session 2-4 ($p = 0.0311$), and Session 3 ($p = 0.0468$). To summarize, the key pattern was that Group B exhibited a significant increase in consonant intelligibility scores from Session 1 to Session 2-1, but Group A reached a significant increase at later sessions such as Session 2-2 or Session 2-3, while Group C showed its first improvement at Session 2-3 or no significant increase at all.

Discussion

This study examined the effect of familiarization on naïve listeners' consonant intelligibility in dysarthric speech. We compared three familiarization conditions (i.e., active vs. passive vs. no explicit familiarization conditions) and found that in all three conditions, naïve listeners' ability to identify consonants improved over time: i.e., the effect of Session was significant and its effect size (the magnitude of difference) was medium to large for all speakers. We further found that the active familiarization method had an advantage over both passive familiarization and no explicit familiarization conditions. The advantage was manifested in terms of two aspects: the mag-

nitude and the rapidness of improvement. First, for all speakers, listeners in the active familiarization condition exhibited higher consonant intelligibility scores compared to the other two conditions in all of the Session 2 rounds. In addition, except for one speaker, we found a significant Group $\times$ Session interaction. The most notable pattern was that consonant intelligibility scores on Session 1 (pre-familiarization) and Session 2-1 (the very first round after familiarization) were significantly different only for Group B. That is, while the other groups reached significant increases at a later session or no improvement at all, listeners in active familiarization significantly improved their ability to identify consonants as early as right after the first familiarization, highlighting the importance of a written transcript inclusion especially for the efficiency of familiarization.

It is worth noting that although the effect was smaller compared to the active and passive conditions, the no-explicit-familiarization group showed improvements in consonant intelligibility for some speakers, similar to the finding by Hustad and Cahill [4] that mere repetitive experience with dysarthric speech might serve as a simple, yet practical way of intelligibility improvement. As in Hustad and Cahill [4], it is undetermined how much of the observed improvement reflects the true learning effect due to familiarization with the speech or the mere practice effect due to familiarization to experimental tasks. The key finding in this study is that the additional effect in passive and active conditions, compared to the no-explicit-familiarization condition, can be attributed to the extra conditions, i.e., the familiarization phases that listeners had in the passive and active conditions.

Regarding the long-term effects of familiarization, we found that familiarization-induced intelligibility enhancement is retained 1 month after familiarization at least to some degree, similar to previous findings in dysarthric speech [2], synthetic speech [20, 37], and nonnative speech [35]. Studies differ in terms of the exact length of a delayed period, ranging from 12 h to 6 months, but findings all support the claim that perceptual learning is not a temporary adjustment but rather a long-lasting effect [3]. It should also be noted that in the current study, the long-term advantage of an active familiarization condition over other conditions was present only for Speaker 4. The active familiarization condition in this study involved more overt experiences of the speech compared to the passive and mere repetition, in that it had a written transcription alongside the audio signal. However, the protocol did not use any explicit method that promotes

listeners to attune to specific phonetic or phonological features [44]. It is possible that more explicit experience of segmental features might yield a maximal learning effect both in the short and long term. More work is needed to investigate what factors influence the longevity of perceptual learning.

Last, several points are worth noting regarding Speaker 3. Listeners showed the least amount of improvement among the speakers examined in this study, and the active familiarization was only as effective as other methods as evidenced by the lack of a significant interaction between Group and Session. Given that perceptual learning is expected to be influenced by acoustic regularities in atypical patterns [3] and/or the nature of acoustic deviances [45], future studies should seek to discover what aspects of the signal (e.g., phonological categories, regularities of acoustic deviances, and the degree of deviances) determine the magnitude of familiarization effects.

This study extends support for segmental benefits of familiarization in dysarthric speech [2, 6, 7] and in other types of atypical speech [18, 35–39]. By basing our analysis on consonant identification tasks rather than word transcription tasks, this study provides conclusive evidence for segmental benefits in perceptual learning in dysarthric speech and supports the previous speculation that use of orthographic transcripts has a beneficial effect on segmental level performance by enabling listeners to adjust the mapping between acoustic-phonetic information and the phonemic representation in the language [6, 39]. From the clinical perspective, the current study highlights the potential of using familiarization as a listener-oriented intervention technique for dysarthria management. Given the consistent finding of familiarization-induced intelligibility improvement in the field, future studies should be directed towards developing an optimal familiarization method in dysarthria. This requires researchers to explore what information could and should be cultivated during familiarization training so that the resulting outcomes will be successful in terms of effectiveness (the amount of improvement), efficiency (rapidness of the improvement), and robustness (stability or longevity of learning). Given the current finding about faster improvement in listeners' consonantal learning in the presence of a written transcript, research must aim to discover what kinds of materials promote phonological learning. An adequate amount of variability is required for perceptual learning to occur, according to the studies in synthesized speech and second-language learning [33, 36], but exactly how to structure the training materials is still under debate. Two main hypotheses,

a traditional minimal-pair approach versus a multiple-cue approach, have been tested in second-language learner studies and phonological disorder studies [36, 46, 47], but they have yet to be studied in the context of perceptual learning in dysarthric speech. Findings will contribute to constructing a standardized format of listener-oriented intervention techniques in dysarthria management.

Acknowledgments

This study was supported by NIDCD R03DC012888A. The author thanks the subjects for their participation and Dr. Torrey Loucks for making a sound booth facility available for our data collection. The author especially thanks Suzanne Nanney for subject recruitments and data collection. Part of this work was presented at the ASHA Annual Convention, Chicago, Ill., USA, November 14–16, 2013 and at the Motor Speech Conference, Sarasota, Fla., USA, February 27–March 2, 2014.

References

- 1 D'Innocenzo J, Tjaden K, Greenman G: Intelligibility in dysarthria: effects of listener familiarity and speaking condition. *Clin Linguist Phon* 2006;20:659–675.
- 2 Borrie SA, McAuliffe MJ, Liss JM, Kirk C, O'Beirne GA, Anderson T: Familiarisation conditions and the mechanisms that underlie improved recognition of dysarthric speech. *Lang Cogn Process* 2012;27:1039–1055.
- 3 Borrie SA, McAuliffe MJ, Liss JM: Perceptual learning of dysarthric speech: a review of experimental studies. *J Speech Lang Hear Res* 2012;55:290–305.
- 4 Hustad KC, Cahill MA: Effects of presentation mode and repeated familiarization on intelligibility of dysarthric speech. *Am J Speech Lang Pathol* 2003;12:198–208.
- 5 Kim H, Nanney S: Familiarization effects on word intelligibility in dysarthric speech. *Folia Phoniatr Logop* 2014;66:258–264.
- 6 Liss JM, Spitzer SM, Caviness JN, Adler C: The effects of familiarization on intelligibility and lexical segmentation in hypokinetic and ataxic dysarthria. *J Acoust Soc Am* 2002;11: 2–6.
- 7 Spitzer SM, Liss JM, Caviness JN, Adler C: An exploration of familiarization effects in the perception of hypokinetic and ataxic dysarthric speech. *J Med Speech Lang Pathol* 2000; 8:285–293.
- 8 Tjaden K, Liss JM: The role of listener familiarity in the perception of dysarthric speech. *Clin Linguist Phon* 1995;9:139–154.
- 9 Tjaden K, Liss JM: The influence of familiarity on judgments of treated speech. *Am J Speech Lang Pathol* 1995;4:39–48.
- 10 Yorkston KM, Beukelman DR: The influence of judge familiarization with the speaker on dysarthric speech intelligibility; in Berry W (ed): *Clinical Dysarthria*. Austin, Pro-Ed, 1983, pp 155–164.
- 11 Garcia JM, Cannito MP: Influence of verbal and nonverbal contexts on the sentence intelligibility of a speaker with dysarthria. *J Speech Hear Res* 1996;39:750–760.
- 12 Bradlow AR, Bent T: Perceptual adaptation to non-native speech. *Cognition* 2008;106:707–729.
- 13 Gass S, Varonis EM: The effect of familiarity on the comprehensibility of nonnative speech. *Lang Learn* 1984;34:65–89.
- 14 Weill SA: Foreign accented speech: encoding and generalization. *J Acoust Soc Am* 2001; 109:2473.
- 15 Ellis LW, Belyukova SA: Effects of training on naïve listeners' judgments of the speech intelligibility of children with severe-to-profound hearing loss. *J Speech Lang Hear Res* 2008;51:1114–1123.
- 16 Loebach JL, Pisoni DB, Svirsky MA: Effects of semantic context and feedback on perceptual learning of speech processed through an acoustic simulation of a cochlear implant. *J Exp Psychol* 2010;36:224–234.
- 17 McGarr NS: The intelligibility of deaf speech to experienced and inexperienced listeners. *J Speech Hear Res* 1983;26:451–458.
- 18 Francis AL, Nusbaum HC, Fenn K: Effects of training on the acoustic phonetic representation of synthetic speech. *J Speech Lang Hear Res* 2007;50:1445–1465.
- 19 McNaughton D, Fallon K, Tod J, Weiner F, Neisworth J: Effect of repeated listening experiences on the intelligibility of synthesized speech. *Augment Altern Commun* 1994;10: 161–168.
- 20 Schwab EC, Nusbaum HC, Pisoni DB: Some effects of training on the perception of synthetic speech. *Hum Factors* 1985;27:395–408.
- 21 Venkatagiri H: Effect of sentence length and exposure on the intelligibility of synthesized speech. *Augment Altern Commun* 1994;10: 96–104.
- 22 Davis MH, Johnsrude IS, Hervais-Adelman A, Taylor K, McGettigan C: Lexical information drives perceptual learning of distorted speech: evidence from the comprehension of noise-vocoded sentences. *J Exp Psychol Gen* 2005;134:222–241.
- 23 Golomb, JD, Peelle JE, Wingfield A: Effects of stimulus variability and adult aging on adaptation to time-compressed speech. *J Acoust Soc Am* 2007;121:1701–1708.
- 24 Pallier C, Sebastian-Galles N, Dupoux E, Christophe A: Perceptual adjustment to time-compressed speech: a cross-linguistic study. *Mem Cognit* 1998;26:844–851.
- 25 Sebastian-Galles N, Dupoux E, Costa A, Mehler J: Adaptation to time-compressed speech: phonological determinants. *Percept Psychophys* 2000;62:834–842.
- 26 Miller N: Measuring up to speech intelligibility. *Int J Lang Commun Disord* 2013;48:601–612.
- 27 Mattys SL: Speech perception; in Daniel R (ed): *The Oxford Handbook of Cognitive Psychology*. Oxford, Oxford University Press, 2011.
- 28 Nygaard LC, Pisoni DB: Talker-specific learning in speech perception. *Percept Psychophys* 1998;60:355–376.
- 29 Yonan CA, Sommers MS: The effects of talker familiarity on spoken word identification in younger and older listeners. *Psychol Aging* 2000;15:88–99.
- 30 Loebach JL, Bent T, Pisoni DB: Multiple routes to the perceptual learning of speech. *J Acoust Soc Am* 2008;124:552–561.
- 31 Borrie SA, McAuliffe MJ, Liss JM, O'Beirne GA, Anderson TJ: The role of linguistic and indexical information in improved recognition of dysarthric speech. *J Acoust Soc Am* 2013;133:474–482.
- 32 Dupoux E, Green K: Perceptual adjustment to highly compressed speech: effects of talker and rate changes. *J Exp Psychol Hum Percept Perform* 1997;23:914–927.
- 33 Greenspan SL, Nusbaum HC, Pisoni DB: Perceptual learning of synthetic speech. *J Exp Psychol Learn Mem Cogn* 1988;14:421–433.
- 34 Pisoni DB, Lively SE, Logan JS: Perceptual learning of nonnative speech contrasts: implications for theories of speech perception; in Nusbaum HC, Goodman J (eds): *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*. Cambridge, Mass., MIT Press, 1994, pp 121–166.
- 35 Nishi K, Kewley-Port D: Training Japanese listeners to perceive American English vowels: influence of training sets. *J Speech Lang Hear Res* 2007;50:1496–1509.

- 36 Nishi K, Kewley-Port D: Nonnative speech perception training using vowel subsets: effects of vowels in sets and order of training. *J Speech Lang Hear Res* 2008;51:1480–1493.
- 37 Eisner F, McQueen JM: The specificity of perceptual learning in speech processing. *Percept Psychophys* 2005;67:224–238.
- 38 Kraljic T, Samuel AG: Generalization in perceptual learning for speech. *Psychon Bull Rev* 2006;13:262–268.
- 39 Norris D, McQueen JM, Cutler A: Perceptual learning in speech. *Cogn Psychol* 2003;47: 204–238.
- 40 Samuel AG, Kraljic T: Perceptual learning for speech. *Atten Percept Psychophys* 2009;71: 1207–1218.
- 41 Kim H, Hasegawa-Johnson M, Perlman A, Gunderson J, Huang T, Watkin K, Frame S: Dysarthric speech database for universal access research. *Proceedings of Inter Speech Conference*, September 22–26, 2008, Brisbane, Australia.
- 42 Boersma P, Weenink D: Praat: doing phonetics by computer (version 5.1.15), 2009. <http://www.praat.org/>.
- 43 Bakeman R: Recommended effect size statistics for repeated measures designs. *Behav Res Methods* 2005;37:379–384.
- 44 Gipson EJ: *Principles of perceptual learning and development*. New York, Prentice Hall College Div, 1969.
- 45 Kim H, Nanne S: Relationship between the spectral characteristics of fricatives and familiarization-induced intelligibility enhancement. *ASHA Annual Convention*, November 20–22, 2014, Orlando, Fla., USA.
- 46 Barlow JA, Gierut JA: Minimal pair approaches to phonological remediation. *Semin Speech Lang* 2002;23:57–67.
- 47 Gierut JA: Enhancement of learning for children with phonological disorders; in Slifka J, Manuel S, Matthies M (eds): *From Sound to Sense: 50+ Years of Discoveries in Speech Communication*. Cambridge, Mass., MIT Press, 2004, pp 164–172.