Published in final edited form as:

Ear Hear. 2009 April ; 30(2): 238–249. doi:10.1097/AUD.0b013e31819769ac.

INTERACTIONS BETWEEN UNSUPERVISED LEARNING AND THE DEGREE OF SPECTRAL MISMATCH ON SHORT-TERM PERCEPTUAL ADAPTATION TO SPECTRALLY-SHIFTED SPEECH

Tianhao Li¹, John J. Galvin III², and Qian-Jie Fu^{1,2}

¹ Biomedical Engineering Department, University of Southern California

² Department of Auditory Implants and Perception, House Ear Institute

Abstract

Objectives—Cochlear implant listeners are able to at least partially adapt to the spectral mismatch associated with the implant device and speech processor via daily exposure and/or explicit training. The overall goal of this study was to investigate interactions between short-term unsupervised learning (i.e., passive adaptation) and the degree of spectral mismatch in normal-hearing listeners' adaptation to spectrally-shifted vowels.

Methods—Normal-hearing subjects were tested while listening to acoustic cochlear implant simulations. Unsupervised learning was measured by testing vowel recognition repeatedly over a five-day period; no feedback or explicit training was provided. In Experiment 1, subjects listened to 8-channel, sine-wave vocoded speech. The spectral envelope was compressed to simulate a 16 mm cochlear implant electrode array. The analysis bands were fixed and the compressed spectral envelope was linearly shifted toward the base by 3.6, 6 or 8.3 mm to simulate different insertion depths of the electrode array, resulting in a slight, moderate, or severe spectral shift. In Experiment 2, half the subjects were exclusively exposed to a severe shift with 8 or 16 channels ("exclusive groups"), and half the subjects were exposed to 8-channel severely-shifted speech, 16-channel severely-shifted speech and 8-channel moderately-shifted speech, alternately presented within each test session ("mixed group"). The region of stimulation in the cochlea was fixed (16 mm in extent, 15 mm from the apex) and the analysis bands were manipulated to create the spectral shift conditions. To determine whether increased spectral resolution would improve adaptation, subjects were exposed to 8- or 16-channel severely-shifted speech.

Results—In Experiment 1, at the end of the adaptation period, there was no significant difference between 8-channel speech that was spectrally-matched or shifted by 3.6 mm. There was a significant, but less-complete adaptation to the 6 mm shift and no adaptation to the 8.3 mm shift. In Experiment 2, for the mixed exposure group, there was significant adaptation to severely-shifted speech with 8 channels, and even greater adaptation with 16 channels. For the exclusive exposure group, there was no significant adaptation to severely-shifted speech, with either 8 or 16 channels.

Conclusions—These findings suggest that listeners are able to passively adapt to spectral shifts up to 6 mm. For spectral shifts beyond 6 mm, some passive adaptation was observed with mixed exposure to a smaller spectral shift, even at the expense of some low frequency information. Mixed exposure to the smaller shift may have enhanced listeners' access to spectral envelope details that were not accessible when listening exclusively to severely-shifted speech. The results suggest that the range of spectral mismatch that can support passive adaptation may be larger than previously

reported. Some amount of passive adaptation may be possible with severely-shifted speech by exposing listeners to a relatively small mismatch in conjunction with the severe mismatch.

Introduction

Most contemporary cochlear implants (CIs) utilize spectrally-based speech processing to restore the tonotopic representation of acoustic information in the auditory system. Typically, an input acoustic signal is divided into a number of frequency analysis bands, and the temporal envelopes extracted from each band are used to modulate pulse trains of current delivered to appropriate electrodes implanted within the cochlea. To restore the normal tonotopic representation, temporal envelopes extracted from low-frequency bands are delivered to apical electrodes, and envelopes extracted from high-frequency bands are delivered to basal electrodes. The acoustic frequency-to-electrode place mapping is critical to the correct transmission of spectral cues. However, the physical relationship of the implanted electrodes to the surviving auditory neurons typically results in some degree of spectral distortion of the acoustic input signal. Longitudinal studies suggest that many CI users are able to at least partially adapt to these new speech patterns, even without explicit training (e.g., Loeb and Kessler, 1995; Spivak and Waltzman, 1990). However, some CI users are unable to adequately adapt, even after years of experience with their device. It is unclear how the degree of spectral distortion may interact with adaptation, especially in an unsupervised learning context.

Spectral distortion may be caused by a spectral shift between the input acoustic frequency and the place of stimulation in the cochlea. Spectral distortion may also be caused by spectral compression of the input acoustic signal onto the limited spatial extent of the electrode array in the cochlea. Typically, the acoustic input is both spectrally-shifted and spectrally-compressed for CI users; for shallow electrode insertion depths, the acoustic input signal may be severely shifted and compressed. Spectral distortion may also be caused by spectral warping, due to “holes” or “dead regions” in the surviving auditory neurons, and/or the proximity of the implanted electrodes to healthy neural populations (Moore, 2001; Shannon et al., 2002a).

There is considerable variability in electrode locations among CI patients. According to *in vivo* computed tomography (CT) studies (Ketten et al., 1998; Skinner et al., 2002), electrode insertion depths among 26 CI patients ranged from 11.9–25.9 mm. In terms of Greenwood's equation (1990), these insertion depths correspond to the characteristic frequencies (CFs) of 3674–308 Hz for the most apical electrode. However, a recent study by Dorman et al. (2007) comparing pitch between acoustic and electric hearing in the same CI patient found the CF of the most apical electrode (insertion depth of 12 mm) to be ~1000 Hz, suggesting that Greenwood's equation may overestimate the CFs of the implanted electrodes. While Greenwood's formula may or may not accurately predict the CFs of implanted electrodes, anecdotal reports suggest that some CI users find that speech sounds “high-pitched” or “squeaky” immediately after initial fitting. This suggests that speech is shifted basally relative to normal hearing, although stimulation rate may also contribute to these initial voice quality percepts. Voices typically begin to sound more natural as CI users gain experience with their implant and speech processor. Thus, while many CI users experience some degree of spectral shift, many are able to at least partially adapt without explicit training. For CI users with a shallow insertion depth, the spectral shift and compression may be too severe to allow full adaptation.

Recognition of spectrally-distorted speech has been extensively researched in acute studies with CI users (e.g., Fu and Shannon, 1999; Baskent and Shannon, 2003, 2004) and/or normal-hearing (NH) subjects listening to acoustic CI simulations (e.g., Shannon et al., 1998; Dorman et al., 1997b). In general, the best performance has been observed with spectrally-matched speech. However, there is some tolerance for spectral mismatch, even when acutely measured.

NH subjects listening to acoustic CI simulations can tolerate ~3 mm of spectral shift with little deficit to speech performance (Fu and Shannon, 1999). Baskent and Shannon (2007) found that certain combinations of spectral shifting and spectral compression/expansion produced better speech recognition than spectral shifting alone, most likely because local frequency-place mismatches were reduced for important speech cues. However, speech understanding declines rapidly as the spectral mismatch increases beyond ~3 mm (Fu and Shannon, 1999).

Perceptual adaptation studies have shown that listeners may tolerate even greater amounts of spectral mismatch and/or spectral warping, given daily exposure or explicit training (Rosen et al., 1999; Fu et al., 2002, 2005; Svirsky et al., 2004; Faulkner et al., 2006; Stacey and Summerfield, 2007). However, the degree of shift may influence the amount and time course of adaptation. Rosen et al. (1999) showed that sentence recognition with upwardly shifted speech (6.5 mm in basilar membrane distance) improved from 1% to 30 % correct after only a few hours of connected discourse training. Dorman and Ketten (2003) found that, after 1 week of daily exposure, one CI user fully adapted to a 3.2 mm basal shift, and partially adapted to a 6.8 mm basal shift. While these previous studies demonstrate adaptation to various amounts of spectral mismatch, there has not been a systematic study of the interaction between perceptual adaptation and degree of spectral shift, especially in an unsupervised learning context.

Potter and Steinberg (1950) asserted that a fixed spatial pattern could be identified equally well when presented to different regions of the cochlea. However, CI speech processors both spectrally shift and compress the acoustic input onto the place of stimulation in the cochlea. For CI users, does difficulty in adaptation arise from the absolute frequency-place mismatch or from distortion to the spectral envelope? If the difficulty arises from distortion to the spectral envelope, then the spectral envelope may be linearly shifted along the cochlea without affecting speech understanding, especially after some amount of perceptual experience. More likely, there may be an optimal tradeoff between distortion to the spectral envelope and the degree of spectral shift that produces the best performance. Smith and Faulkner (2006) found that NH listeners' adaptation to a simulated 'hole' in the spectrum (i.e., spectral warping) was similar to adaptation to spectral shift. This suggests that adaptation to spectral envelope distortion may be evaluated in terms of local frequency mismatches. Most previous perceptual adaptation studies have incorporated two shift conditions (at most) and explicit training, making it difficult to observe the complex interactions between perceptual adaptation, spectral shifting and spectral envelope distortion. The present study examined these questions by delivering a fixed spectral envelope to three locations along the cochlea to observe interactions between perceptual adaptation and spectral shifting.

Explicit training can improve adaptation to a spectral shift, assuming that sounds can be discriminated. It is unclear whether the neural system adapts to spectrally-shifted speech according to speech patterns held in long-term memory (formed during normal hearing or hearing-impaired experience), or whether the neural system forms new patterns via explicit training. Li and Fu (2007) suggested that explicit training is necessary for adaptation to severely-shifted speech, as non-lexical label training did not generalize to recognition of severely-shifted vowels, at least within a five-day study period. Rather than an explicit training protocol, a "test-only" protocol was utilized in the present study, in which subjects were repeatedly tested over a five-day study period. The test-only protocol was used to ensure that perceptual adaptation was modulated by speech patterns in long-term memory. While explicit training may accelerate the learning process, a test-only protocol may provide greater insight into speech patterns held in the central nervous system, and their role in perceptual adaptation.

Gradual exposure to a severe spectral shift may improve listeners' overall adaptation, as well as reduce the stress of learning new speech patterns. Fu et al. (2002) found partial adaptation

in CI users after three months of continuous exposure to a one-octave shift; no explicit training was provided. When the same shift was gradually introduced over an 18-month period, Fu and Galvin (2007) found better adaptation. Similarly, Svirsky et al. (2003) found that gradual exposure accelerated NH listeners' adaptation to a 6.5 mm basal shift. It is possible that by successively introducing spectral shifts within the passive adaptation range, listeners may "bridge" large spectral mismatches without explicit training or feedback. In Experiment 2, a "mixed exposure" protocol was used to determine whether mixed exposure to moderately and severely shifted speech would improve recognition of severely shifted speech.

Because spectral cues have been shown to be the most critical for vowel recognition (van Schijndel et al., 2001; Shannon, 2002b; Turner et al., 1995), the present study investigated the effects of spectral shifting on listeners' vowel recognition performance. In Experiment 1, vowel recognition was compared for different degrees of spectral shift, using 8-channel, sine-wave vocoded speech. The spectral envelope was spectrally compressed (to simulate a fixed electrode array extent), then delivered to different cochlear locations (to simulate different electrode insertion depths), resulting in a slight, moderate or severe basal shift. In Experiment 2, the spectral envelope was spectrally compressed and severely shifted. Adaptation to the severe shift was measured with and without exposure to 8-channel moderately-shifted speech. To examine listeners' sensitivity to spectral envelope resolution, adaptation to the severe shift was measured with 8 or 16 channels.

EXPERIMENT 1: Effect of spectral shifting on perceptual adaptation

Methods

A. Subjects—Eight NH subjects participated in the study (aged 18 to 35 years old, 4 men and 4 women); all subjects were native speakers of American English. All subjects had pure-tone thresholds better than 20 dB HL at octave frequencies from 125 Hz to 8000 Hz. Subjects had no prior experience with acoustic CI simulations before the study. All subjects were paid for their participation. Note that data for the 8.3 mm severe shift condition are from Fu et al. (2005); four NH subjects participated in that study.

B. Speech materials—Speech stimuli were 12 medial vowels presented in a /h-V-d/ context (i.e., "had," "hod," "hawed," "head," "heard," "hid," "heed," "hood," "hud," "who," "hayed," "hoed"). Vowel tokens were digitized natural productions from five male and five female talkers, randomly drawn from the speech samples recorded by Hillenbrand et al. (1995). Across all 120 tokens (12 vowels*10 talkers), the mean fundamental frequency (F0) was 178 Hz ($\pm$ 55 Hz), the mean first formant frequency (F1) was 570 Hz ($\pm$ 143 Hz), and the mean second formant frequency (F2) was 1625 Hz ($\pm$ 528 Hz).

C. Signal processing and spectral shift conditions—An 8-channel sine-wave vocoder was used to simulate CI speech processing, which was implemented as follows. The input speech signal was filtered into 8 frequency analysis bands (4th-order Butterworth filters). The temporal envelope was extracted from each analysis band by half-wave rectification and low-pass filtering (4th-order Butterworth filter with corner frequency at 160 Hz). The temporal envelope from each channel was used to modulate a corresponding sine-wave carrier; the sine-wave carrier frequencies were varied according to the experimental condition. The modulated sine-waves were summed and the output was adjusted to have the same long-term root-mean-square (RMS) energy as the input speech signal (65 dB). Sine-wave vocoders, rather than noise-band vocoders, were used to restrict the place of stimulation for each channel within the cochlea, as noise carrier bands would excite a broader region for each channel. Note that Dorman et al. (1997a) found no difference in performance between sine-wave and noise-band vocoders for speech recognition in quiet.

Three conditions of upward shift between the frequency analysis bands and sine-wave carriers were compared. Relative to the center frequency of the most apical analysis band, the most apical carrier was shifted by 3.6 mm (slight shift), 6 mm (moderate shift), or 8.3 mm (severe shift). Thus, the degree of spectral shift was linearly increased in terms of cochlear distance. Table 1 shows the distribution of the analysis and carrier bands for three shift conditions. The overall input frequency range was fixed for the three shift conditions (200–7000 Hz). The spatial distribution of center frequencies was calculated according to Greenwood's (1990) formula, assuming a 35-mm long cochlea. Center frequencies of analysis bands were separated by 2.7 mm. The output carrier bands were upwardly shifted to simulate different insertion depths of a 16-mm-long, 8-electrode array with 2-mm electrode spacing. For the 3.6 mm shift, the overall output frequency range was 455–5332 Hz; the degree of spectral mismatch (in terms of cochlear distance) gradually decreased from 3.6 mm for the most apical channel to –1.6 mm for the most basal channel. For the 6 mm shift, the overall output frequency range was 683–7416 Hz; the degree of spectral mismatch gradually decreased from 6 mm for the most apical channel to 0.8 mm for the most basal channel. For the 8.3 mm shift, the overall output frequency range was 999–10290 Hz; the degree of spectral mismatch gradually decreased from 8.3 mm for the most apical channel to 3.1 mm for the most basal channel.

Figure 1 illustrates the three spectral shift conditions. Note that while the spectral envelope from the analysis frequency range was compressed onto the carrier frequency ranges, the compression was uniform across the three shift conditions. Thus, the spatial patterns within the cochlea were similar across conditions, but delivered to different cochlear regions. Also, within each shift condition, there were different degrees of spectral mismatch for different frequency regions of the acoustic input due to spectral compression.

D. Test and adaptation protocol—For all conditions, vowel recognition was measured using a 12-alternative forced choice paradigm. A stimulus was randomly selected (without replacement) from the stimulus set and presented to the subject. The subject responded by clicking on one of 12 response boxes, after which a new stimulus was presented. The response boxes were labeled with an /h-vowel-d/ word (i.e., “had”, “hod”, “head” etc). No training or feedback was provided.

Subjects were randomly assigned to 2 different groups (4 subjects in each group). The first group received five days of repeated testing with vowels shifted by 3.6 mm, and the second group received five days of repeated testing with vowels shifted by 6 mm. The data for the 8.3 mm shift condition were previously reported in Fu et al. (2005), and were collected in an identical manner to the present study. On Day 1, baseline vowel recognition performance was measured with unprocessed speech, as well as with 8-channel, spectrally-matched, sine-wave vocoded speech (0 mm shift, i.e., the center frequencies of the analysis and carrier bands were matched). Baseline performance measures were also used to familiarize subjects with the test procedure and the vocoder processing. On Day 5, baseline performance was re-measured with unprocessed speech and with 8-channel, spectrally-matched speech. These follow-up measures were conducted to observe any incidental learning of the vowel stimuli or vocoder processing while adapting to 8-channel, spectrally-shifted speech. This was deemed important because Davis et al. (2005) observed rapid adaptation to 6-channel, spectrally-matched sentences. The adaptation protocol consisted of five consecutive days of repeated testing with spectrally-shifted speech. Three to four test blocks were administered during each session, on each of Days 1–5; each session lasted ~45 minutes. In each of the test blocks, subjects were tested using the entire stimulus set (120 tokens: 12 vowels * 10 talkers). Again, no preview, feedback, or other explicit training was provided.

Results

Figure 2 shows recognition performance for 8-channel, spectrally-matched speech and for the three spectral shift conditions, as a function of test session. Table 2 shows the results of one-way, repeated-measures (RM) ANOVAs within each processing condition (with test session as factor); significant differences are shown in bold type. Mean baseline vowel recognition was 91% correct with unprocessed speech. Mean performance with 8-channel, sine-wave vocoded, spectrally-matched speech was 78% correct on Day 1 and 82% correct on Day 5. For the 3.6 mm shift, mean performance was 66% correct on Day 1 (~12 percentage points lower than baseline measures with 8-channel, spectrally-matched speech), and 81% correct on Day 5. For the 6 mm shift, mean performance was 41% correct on Day 1 (~36 percentage points lower than baseline measures) and 68% correct on Day 5. For the 8.3 mm shift, mean performance was only 8% correct on Day 1 (near chance level of 8.33 percent correct) and 11% correct on Day 5. As shown in Table 2, vowel recognition performance significantly improved over the five-day study period for the 3.6 and 6 mm shifts; the small 4 point improvement for spectrally-matched speech was also found to be significant. Bonferroni t-tests confirmed the significant difference in post-adaptation performance between the 3.6 and 6 mm shift conditions ($p=0.005$), and between the 6 and 8.3 mm shift conditions ($p<0.001$). However, there was no significant difference in post-adaptation performance between spectrally-matched speech and the 3.6 mm shift condition ($p=0.25$).

The percent of information transmitted for first and second formants (F1 and F2) was calculated using methods introduced by Miller and Nicely (1955). Table 3 shows the F1 and F2 feature categories for the 12 vowels. Figure 3 shows the mean percent of F1 and F2 information transmitted on Day 1 (black bars) and Day 5 (white bars) for the different speech processing conditions. Table 2 shows the results of one-way RM ANOVAs within each shift condition (comparison between Day 1 and Day 5) for the percent of F1 and F2 information transmitted; significant differences are shown in bold type. For spectrally-matched speech, (Panel A), the percent of F2 information transmitted improved over the five-day study period. For the 3.6 mm (Panel B) and 6 mm (Panel C) shift conditions, the percent of F1 and F2 information transmitted improved over the five-day study period. While vowel recognition performance was poorer with the 6 mm shift than with the 3.6 mm shift, the relative improvement in percent of F2 information transmitted was greater for the 6 mm shift. For the 8.3 mm shift, (Panel D), there was little change in the percent of F1 or F2 information transmitted over the five-day study period. As shown in Table 2, there was a significant improvement in F1 information transmitted for the 3.6 mm and 6 mm shift conditions, and in F2 information transmitted for the spectrally-matched, 3.6 mm and 6 mm conditions.

Experiment 2: Effect of spectral resolution on passive adaptation to severely-shifted vowels measured in the presence or absence of mixed exposure to moderately-shifted vowels

The data from Experiment 1 showed that severe spectral shift greatly limited the degree of adaptation. Specifically, there was no adaptation to the 8.3 mm shift. It is possible that increasing the spectral resolution may improve adaptation, as more acoustic information is usually helpful and/or necessary under noisy and difficult listening conditions. Also, gradual exposure to a severe shift may provide better adaptation, even in an unsupervised learning context. For CI users, gradual exposure may necessitate some tradeoffs between spectral mismatch and the bandwidth of acoustic information mapped onto the electrodes. As users gradually accommodate the spectral mismatch, more acoustic information can be included. However, previous studies showed no significant benefit in post-training performance for gradual adaptation. Svirskey et al. (2003) showed that while gradual exposure accelerated adaptation to a 6.5 mm basalward shift, post-training performance was not significantly better

than when the shift was abruptly introduced. Similarly, Faulkner et al. (2005) showed no significant benefit for gradual adaptation. In the present study, a mixed exposure protocol was used to provide some control over the experimental conditions, as well as to test potential implementation of such a protocol (e.g., loading different maps with different acoustic input ranges). For all experimental conditions, the place of stimulation in the cochlea was fixed to simulate a shallow electrode insertion depth. Rather than gradually introducing the severe shift or the 8-channel processing, subjects were alternately exposed to 8-channel moderately-shifted speech, 8-channel severely-shifted speech, and 16-channel severely-shifted speech; the speech processing conditions were tested in random order during each session.

Methods

A. Subjects—Eight NH subjects (aged from 18 to 23 years old; 6 males and 2 females) participated in the study; all subjects were native speakers of American English. All subjects had pure-tone thresholds better than 20 dB HL at octave frequencies from 125 Hz to 8000 Hz. None of these subjects participated in Experiment 1, and none had any experience with acoustic CI simulations before the study. All subjects were paid for their participation.

B. Speech materials and signal processing—Speech materials (vowels) and basic signal processing (i.e., sine-wave vocoders) were the same as in Experiment 1, except for the number of channels (8 or 16) and the analysis band/sine-wave carrier parameters.

C. Spectral shift conditions—Two adaptation conditions were studied: 1) exposure to only 8- or 16-channel severely-shifted speech (“exclusive exposure”) and 2) exposure to 8-channel severely-shifted speech, 16-channel severely-shifted speech and 8-channel moderately-shifted speech (“mixed exposure”). Table 4 shows the corner and center frequencies for each experimental condition. For all spectral shift conditions, the overall output frequency range was fixed (999–10290 Hz) to simulate a shallow insertion depth. The spatial distribution was calculated according to Greenwood's (1990) formula, assuming a 35-mm long cochlea. The input frequency range was manipulated to create the different spectral shift conditions. For severely-shifted speech, the overall input frequency range was 200–7000 Hz; the acoustic input was spectrally compressed onto the output frequency range. As shown in Table 4, there was a slight difference in the spectral shift of the most apical channel between the 8-channel (8.3 mm) and 16-channel (8.5 mm) processing. The frequency analysis range for the most apical channel for the 8-channel processor was 200–359 Hz. This range was divided between the two most apical channels for the 16-channel processor. To maintain a similar degree of spectral shift for the 8- and 16-channel processors, the most apical channel for the 16-channel processor was shifted by 8.5 mm and the second-most apical channel was shifted by 8.1 mm, resulting in an average shift of 8.3 mm. For moderately-shifted speech, the overall input frequency range was 455–12275 Hz; the input frequency range was high-pass filtered and spectrally compressed to reduce the spectral mismatch between the most apical analysis and carrier bands; thus, there was some loss of low-frequency information. The most apical channel was shifted by 4.3 mm. Note that the “moderate” condition in Experiment 2 was different from that in Experiment 1 (i.e., different overall acoustic input frequency range, less spectral mismatch for the most apical channel). This manipulation represents a potentially viable option for gradually introducing a severe spectral shift to CI users with a shallow electrode insertion depth. Figure 4 illustrates the three experimental conditions.

D. Test and adaptation protocol—Adaptation to spectrally-shifted vowels was measured as in Experiment 1, i.e., repeated testing (without explicit training or feedback) over the five-day study period. Subjects were randomly assigned into one of two groups (4 subjects in each group). In the first group (“exclusive exposure”), subjects listened exclusively to 16-channel severely-shifted speech. For comparison purposes, data from Fu et al. (2005) were used to

show performance with exclusive exposure to 8-channel severely-shifted speech. In the second group (“mixed exposure”), subjects listened to 8- and 16-channel severely-shifted speech, as well as 8-channel moderately-shifted speech. Table 5 shows the timeline for Experiment 2. For both test groups, baseline performance was first measured with unprocessed speech as well as 8- and 16-channel spectrally-matched speech (0 mm shift); baseline performance was re-measured after the five-day adaptation period. In the exclusive exposure protocol, subjects were repeatedly tested over five consecutive days, exclusively listening to 16-channel severely-shifted speech (or 8-channel severely-shifted speech in Fu et al., 2005). Three to four test blocks were administered during each session, on each of Days 1–5; each session lasted ~45 minutes. In each of the test blocks, subjects were tested using the entire stimulus set (120 tokens: 12 vowels * 10 talkers). In the mixed exposure protocol, subjects were repeatedly tested for five days while alternately listening to 8- and 16-channel severely-shifted speech and 8-channel moderately-shifted speech. Within each session, subjects completed 3–4 test blocks for each of the speech processing conditions. During each test block, subjects were tested using the entire stimulus set (120 tokens: 12 vowels * 10 talkers). Each session lasted ~2.5 hours. The test order of experimental conditions was randomized across sessions and subjects.

Results

Figure 5 shows mean performance for the different spectral shift conditions and adaptation protocols. Table 6 shows the results of one-way RM ANOVAs within each adaptation protocol and speech processing condition (with test session as factor). Significant differences are shown in bold type. There was no improvement in recognition of 8- or 16-channel severely shifted speech with the exclusive exposure protocol. Mean performance with 16 channels was 11% correct on Day 1 and Day 5 (nearly chance level of 8.33 percent correct). Mean performance with 8 channels was 8% correct on Day 1 and 11% correct on Day 5. With the mixed exposure protocol, mean performance improved from Day 1 to Day 5 for all three speech processing conditions. Mean performance with 8-channel moderately-shifted speech improved from 54% to 67% correct. Mean performance with 8-channel severely-shifted speech improved from 13% to 21% correct. Mean performance with 16-channel severely-shifted speech improved from 17% to 34% correct. As shown in Table 6, vowel recognition performance significantly improved for all processing conditions with the mixed exposure protocol, but not with the exclusive exposure protocol. A two-way ANOVA showed a significant effect for adaptation protocol for both 8-channel [$F(1,30) = 13.768$, $p < 0.001$] and 16-channel speech [$F(1,30) = 68.996$, $p < 0.001$].

Figure 6 shows the mean percent correct of F1 and F2 information transmitted on Day 1 (black bars) and Day 5 (white bars) for the various speech processing conditions; the top panels show results for the exclusive exposure group and the bottom panels show results for the mixed exposure group. Table 6 shows the results of one-way RM ANOVAs comparing the percent of F1 and F2 information transmitted between Day 1 and Day 5, for the various speech processing conditions; significant differences are shown in bold type. For the exclusive exposure protocol, there was little difference in F1 and F2 information transmitted from Day 1 to Day 5, for 8-channel (Panel A) or 16-channel (Panel B) severely-shifted speech. For 8-channel moderately-shifted speech (Panel C) and 16-channel severely-shifted speech (Panel E), the percent of F1 and F2 information transmitted improved from Day 1 to Day 5. With 8-channel severely-shifted speech (Panel D), there was little change in the percent of F1 and F2 information transmitted. As shown in Table 6, there was no significant change in the percent of F1 or F2 information transmitted for the exclusive exposure group. For the mixed exposure group, the percent of F2 information transmitted significantly improved for 8-channel moderately-shifted speech and 16-channel severely-shifted speech; there was no significant change in the percent of F1 or F2 information transmitted for 8-channel severely-shifted speech.

Discussion

In the present study, passive adaptation to spectrally-shifted vowels was measured over a five-day period using a test-only protocol. In general, acute performance and adaptation were significantly affected by the degree of spectral shift. The present study provides several new insights regarding listeners' adaptation to a spectral mismatch. First, while previous studies have shown adaptation with explicit training or daily exposure, significant learning was observed in the present study with unsupervised adaptation. Second, there seems to be a non-linear effect for the degree of spectral shift on passive adaptation. Delivering the same spectral envelope to different cochlear locations had a strong effect on performance, even after short-term adaptation. Third, introduction of a severe shift in conjunction with a moderate shift provided better passive adaptation and greater sensitivity to spectral envelope details.

While acute studies have shown that spectrally-matched speech processors are optimal for CI users, perceptual learning studies have shown that both CI and NH listeners are able to adapt to spectrally-shifted speech. Most of these previous learning studies have utilized some form of explicit training, feedback and/or stimulus preview, and suggest that listeners may adapt by creating new central speech pattern templates via training. The present study utilized a test-only protocol, in which adaptation was measured without training, feedback, or stimulus preview; therefore, adaptation presumably occurred in relation to the central speech patterns developed during normal hearing.

In the present study, there was an interaction between unsupervised learning and the degree of spectral shift. With exclusive exposure to slight and moderate shifts, significant amounts of passive adaptation were observed. There was no passive adaptation with exclusive exposure to a severe shift. However, when the severe shift was introduced via the mixed exposure protocol, significant passive adaptation was observed. The mixed exposure protocol involved alternating exposure to the moderate and severe shifts, as well as to 8- and 16-channel processing. In a true gradual exposure protocol, listeners would be incrementally exposed to increasing amounts of spectral shift (e.g., Fu and Galvin, 2007). With the mixed exposure protocol, the degree of passive adaptation of a severe shift (from chance level to 34% correct) was similar to that reported in previous studies with explicit training. However, the materials and methods used for testing and training may have influenced training outcomes. Rosen et al. (1999) showed that sentence recognition with a 6.5 mm basalward shift improved from 1% to 30% correct with connected discourse training. Vowel recognition improved by ~20 percentage points, slightly less than the improvement observed with the 6 mm shift in Experiment 1 of the present study (in which subjects received no explicit training). Fu et al. (2005) reported that with phonetic contrast training, recognition of severely-shifted (8.3 mm basal shift) vowels improved from chance level to 29% correct; with sentence training (modified connected discourse), severely-shifted vowel recognition improved from chance level to 16% correct. Stacey and Summerfield (2008) also explored the effects of training methods and materials for understanding of spectrally-shifted speech (8-channel, 6 mm shift, similar to Experiment 1 in the present study). In contrast to Fu et al. (2005), they found better improvements in word, sentence and vowel recognition when training with words and sentences; phonetic contrast discrimination training provided the least benefit. The phonetic contrast methodology and materials were quite different between the two studies: discrimination along a continuum between two synthesized phonemes (e.g., "a" vs. "uh") in Stacey and Summerfield (2008) vs. identification of targeted medial contrasts using words (e.g., "bad" vs. "bud") in Fu et al. (2005). The improved vowel recognition for the 6 mm shift in Experiment 1 (without any explicit training) was greater than that observed for any of the explicit training methods or tests used in Stacey and Summerfield (2008). Nevertheless, explicit training may be necessary to improve performance beyond that which can be passively learned. Li and Fu (2007) found that while non-lexical label training improved discrimination of

severely-shifted vowels, the training did not generalize to identification of severely-shifted vowels. It remains unclear whether the improved vowel recognition observed in the present study would provide better sentence and/or word recognition. The present data points to the difficulty in disentangling the benefits of explicit training from those of passive learning. Depending on the degree of spectral shift and/or how the shift is introduced, significant amounts of passive adaptation are possible, and the improvements in performance may be comparable to those observed with explicit training, under certain conditions.

In the present study, passive adaptation was measured over a five-day study period. It is possible that severely-shifted speech requires a longer adaptation period. However, asymptotic performance was achieved for most test conditions within five days. In addition, Fu et al. (2002) found that, in a three-month study with CI subjects who continuously wore severely-shifted speech processors, most of the major improvements in performance occurred within the first week of adaptation, suggesting that the initial adaptation process may occur fairly rapidly. The longer-term, more slowly-evolving adaptation that follows may require more complex neural plasticity and cognitive activity. Adaptation in CI simulation studies may be different from real CI users' adaptation. An otherwise healthy auditory system adapting to spectrally-shifted speech may be able to access speech cues that are less available to the impaired/implanted auditory system.

While previous studies have examined adaptation to spectral shift using one or two shift conditions (e.g., Rosen et al., 1999; Fu et al., 2002; Dorman and Ketten, 2003; Svirsky et al., 2004), the present study examined interactions between spectral shifting and adaptation for several shift conditions. The data from Experiment 1 showed that the decline in speech performance was less severe for the 3.6 mm and 6 mm shifts than for the 8.3 mm shift. For the 3.6 mm and 6 mm shifts, mean performance improved after five days of repeated testing by 15 and 27 percentage points, respectively; performance with the 8.3 mm shift did not significantly improve and remained near chance level. With exclusive exposure, performance remained good for spectral shifts ranging from 0 – 6 mm, beyond which performance dropped precipitously. In Experiment 2, mixed exposure to the moderate 4.3 mm shift softened the sharp decline in performance with the 8.3 mm severe shift. The data from Experiment 1 further suggests that the degree of spectral shift dominates passive adaptation to spectrally-shifted speech. The same spectral envelope was delivered to three cochlear locations, and post-adaptation performance decreased nonlinearly as the degree of spectral shift was linearly increased.

It is interesting to consider “what” was learned, in terms of formant frequency cues. Figure 3 shows that for spectrally-matched speech and for the 6 mm shift, there was a greater improvement in the percent of F2 information transmitted than for F1 information transmitted. Similarly, for the 8-channel moderate shift in Experiment 2 (see Panel C of Figure 6), there was a greater improvement in the percent of F2 information transmitted than for F1 information transmitted. For both Experiments 1 and 2 with the 8.3 mm shift, reception of F2 cues was not markedly better than that for F1 cues. In general, for shifts up to 6 mm, listeners seemed to more strongly adapt to mismatched F2 information. However, there is some evidence that listeners may have adapted to shifted F1 cues. For acutely measured performance with the 3.6 mm shift, confusion matrix analyses showed that ‘hid,’ ‘head,’ ‘hud,’ and ‘who’d’ were most often confused with ‘head,’ ‘had,’ ‘hod,’ and ‘hoed’ (all of which have relatively higher F1 values), respectively. After five days of repeated testing, most of these vowels could be correctly identified.

The relative importance of F1 and F2 cues for vowel recognition may also play a role in understanding the nature of adaptation in the present study. Improved reception of F2 cues may have contributed more strongly to better overall performance simply because F2 cues are more

important to vowel recognition in quiet. However, listening conditions may also influence the relative importance of formant information and, in turn, the nature of perceptual adaptation. Parikh and Loizou (2005) showed that in noisy listening conditions, listeners relied on accurate F1 frequency information along with partial F2 envelope information for vowel identification. The neural system may make use of different acoustic cues for speech identification under different listening conditions. In the present study, the improvements in overall performance and reception of F2 cues for vowel recognition were measured in quiet. It is unclear whether these improvements would hold for other speech tests and/or other listening conditions. Faulkner et al. (2005) showed that the results of phonetic training did not necessarily generalize to recognition of connected speech. In the present study, word and/or sentence recognition was not tested, so it is difficult to know whether listeners would have received these “real-world” benefits.

Consistent with previous studies (e.g., Svirsky et al., 2003; Fu and Galvin, 2007), the results of the present study provide evidence that gradual or mixed exposure, rather than abrupt exposure to a severe shift, may be more beneficial to adaptation. In the present study, a test-only protocol was used, and the degree of spectral shift was larger than that used in previous studies. In terms of the nature of perceptual learning with the mixed protocol, it is unclear whether listeners developed new stimulation patterns for severely-shifted speech in relation to moderately-shifted speech (which were presumably within the acceptable range of mismatch, relative to the normal patterns). With explicit training, listeners may use feedback to create new central patterns. With passive learning, listeners are presumed to adapt in relation to the previously-learned central patterns. Interestingly, the mixed exposure protocol allowed listeners to access spectral envelope details that were inaccessible with the exclusive exposure protocol. It is unclear why these details were less accessible in the exclusive exposure condition. While the source of improvement is unclear, the mixed exposure protocol could be implemented sequentially in CI patients with shallow electrode insertion depth, gradually mapping more low frequency information onto the electrode array. Different maps with different degrees of spectral mismatch could be loaded onto CI patients’ speech processors. Similar to the present study, CI users could alternate between these maps and listen to the different spectral shifts and different amounts of speech information contained in each map. Successful accommodation may ultimately depend on CI users’ willingness to listen to the more distorted speech pattern, given the opportunity to listen to the less distorted pattern. Again, explicit training may help to accelerate and improve adaptation to severely shifted speech, compared to a passive gradual or mixed adaptation protocol.

Conclusions

NH listeners’ short-term adaptation to different degrees of spectral shift was measured over a five-day study period. No feedback or explicit training was provided.

The results show that:

1. Given a short period of unsupervised adaptation, NH listeners’ vowel recognition was largely unaffected by upward spectral shifts of up to 6 mm. Vowel recognition declined sharply for an 8.3 mm basalward shift.
2. There was a non-linear interaction between the degree of spectral shift and passive adaptation. Compared to spectral envelope distortion, spectral shifting had a stronger effect on speech performance.
3. Alternating exposure to a severe and moderate spectral shift provided better passive adaptation to a severe shift and greater sensitivity to spectral envelope details.

These data suggest that there may be an optimum range (0 – 6 mm) of spectral mismatch within which some passive adaptation is possible; this range may be larger if listeners are alternately exposed to moderate spectral shifts, even at the expense of some speech information.

Acknowledgments

The authors wish to thank Dr. Gail Donaldson and three anonymous reviewers for insightful comments and detailed suggestions on this manuscript. This work was supported by NIH/NIDCD 004792.

References

1. Baskent D, Shannon RV. Speech recognition under conditions of frequency-place compression and expansion. *J. Acoust. Soc. Am* 2003;113:2064–2076. [PubMed: 12703717]
2. Baskent D, Shannon RV. Frequency-place compression and expansion in cochlear implant listeners. *J. Acoust. Soc. Am* 2004;116:3130–3140. [PubMed: 15603158]
3. Baskent D, Shannon RV. Combined effects of frequency compression-expansion and shift on speech recognition. *Ear Hear* 2007;28:277–289. [PubMed: 17485977]
4. Davis MH, Johnsrude IS, Hervais-Adelman A, et al. Lexical information drives perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences. *J. Exp. Psychol* 2005;134:222–241.
5. Dorman MF, Loizou PC, Rainey D. Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs. *J. Acoust. Soc. Am* 1997a; 102:2403–2411. [PubMed: 9348698]
6. Dorman MF, Loizou PC, Rainey D. Simulating the effect of cochlear-implant electrode insertion-depth on speech understanding. *J. Acoust. Soc. Am* 1997b;102:2993–2996. [PubMed: 9373986]
7. Dorman MF, Ketten D. Adaptation by a cochlear-implant patient to upward shifts in the frequency representation of speech. *Ear Hear* 2003;24:457–460. [PubMed: 14534415]
8. Dorman MF, Spahr T, Gifford R, et al. An electric frequency-to-place map for a cochlear implant patient with hearing in the nonimplanted ear. *J. Assoc. Res. Otolaryngol* 2007;8:234–240. [PubMed: 17351713]
9. Faulkner, A.; Rosen, S.; Jackson, A. Relative effectiveness of training methods for adaptation to spectrally-shifted speech.. Proceedings of the 2005 Mid-Winter Meeting of Assoc. Res. Otolaryngol; 2005. Poster #382
10. Faulkner A, Rosen S, Norman C. The right information may matter more than frequency-place alignment: simulations of frequency-aligned and upward shifting cochlear implant processors for a shallow electrode array insertion. *Ear Hear* 2006;27:139–152. [PubMed: 16518142]
11. Fu Q-J, Shannon RV. Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing. *J. Acoust. Soc. Am* 1999;105:1889–1900. [PubMed: 10089611]
12. Fu QJ, Shannon RV, Galvin JJ. Perceptual learning following change in the frequency-to-electrode assignment with the Nucleus-22 cochlear implant. *J. Acoust. Soc. Am* 2002;112:1664–1674. [PubMed: 12398471]
13. Fu QJ, Galvin JJ. The effects of short-term training for spectrally mismatched noise-band speech. *J. Acoust. Soc. Am* 2003;113:1065–1072. [PubMed: 12597199]
14. Fu QJ, Galvin JJ. Perceptual learning and auditory training in cochlear implant recipients. *Trends. Amplif* 2007;11:193–205. [PubMed: 17709574]
15. Fu QJ, Nogaki G, Galvin JJ. Auditory training with spectrally shifted speech: implications for cochlear implant patient auditory rehabilitation. *J. Assoc. Res. Otolaryngol* 2005;6:180–189. [PubMed: 15952053]
16. Greenwood DD. A cochlear frequency-position function for several species – 29 years later. *J. Acoust. Soc. Am* 1990;87:2592–2605. [PubMed: 2373794]
17. Hillenbrand J, Getty LA, Clark MJ, et al. Acoustic characteristics of American English vowels. *J. Acoust. Soc. Am* 1995;97:3099–3111. [PubMed: 7759650]
18. Ketten DR, Vannier MW, Skinner MW, et al. In vivo measures of cochlear length and insertion depth of nucleus cochlear implant electrode arrays. *Ann. Otol. Rhinol. Laryngol* 1998;175:1–16.

19. Li T, Fu QJ. Perceptual adaptation to spectrally-shifted vowels: training with non-lexical labels. *J. Assoc. Res. Otolaryngol* 2007;8:32–41. [PubMed: 17131213]
20. Loeb GE, Kessler DK. Speech recognition performance over time with the Clarion cochlear prosthesis. *Ann. Otol. Rhinol. Laryngol. Suppl* 1995;166:290–292. [PubMed: 7668675]
21. Moore BCJ. Dead regions in the cochlea: diagnosis, perceptual consequences, and implications for the fitting of hearing aids. *Trends Amplif* 2001;5:1–34.
22. Parikh G, Loizou PC. The influence of noise on vowel and consonant cues. *J. Acoust. Soc. Am* 2005;118:3874–3888. [PubMed: 16419830]
23. Potter R, Steinberg J. Toward the specification of speech. *J. Acoust. Soc. Am* 1950;22:807–820.
24. Rosen S, Faulkner A, Wilkinson L. Adaptation by normal listeners to upward spectral shifts of speech: implications for cochlear implants. *J. Acoust. Soc. Am* 1999;106:3629–3636. [PubMed: 10615701]
25. Shannon RV, Zeng F-G, Wygonski J. Speech recognition with altered spectral distribution of envelope cues. *J. Acoust. Soc. Am* 1998;104:2467–2476. [PubMed: 10491708]
26. Shannon RV, Galvin JJ 3rd, Baskent D. Holes in hearing. *J. Assoc. Res. Otolaryngol* 2002a;3:185–199. [PubMed: 12162368]
27. Shannon RV. The relative importance of amplitude, temporal, and spectral cues for cochlear implant processor design. *Am. J. Audiol* 2002b;11:124–127. [PubMed: 12691223]
28. Skinner MV, Ketten DR, Holden LK, et al. CT-derived estimation of cochlear morphology and electrode array position in relation to word recognition in Nucleus-22 recipients. *J. Assoc. Res. Otolaryngol* 2002;3:332–350. [PubMed: 12382107]
29. Smith MW, Faulkner A. Perceptual adaptation by normally hearing listeners to a simulated “hole” in hearing. *J. Acoust. Soc. Am* 2006;120:4019–4030. [PubMed: 17225428]
30. Spivak LG, Waltzman SB. Performance of cochlear implant patients as a function of time. *J. Speech Hear. Res* 1990;33:511–519. [PubMed: 2232769]
31. Stacey PC, Summerfield AO. Effectiveness of computer-based auditory training in improving the perception of noise-vocoded speech. *J. Acoust. Soc. Am* 2007;121:2923–2935. [PubMed: 17550190]
32. Stacey PC, Summerfield AO. Comparison of word-, sentence-, and phoneme-based training strategies in improving the perception of spectrally distorted speech. *J. Speech Lang. Hear. Res* 2008;51:526–538. [PubMed: 18367694]
33. Svirsky MA, Silveira A, Neuburger H, et al. Long-term auditory adaptation to a modified peripheral frequency map. *Acta Otolaryngol* 2004;124:381–386. [PubMed: 15224858]
34. Svirsky, MA.; Sinha, S.; Neuberger, H., et al. Gradual adaptation to shifts in the peripheral acoustic frequency map.. *Proceedings of the 2003 Mid-Winter Meeting of Assoc. Res. Otolaryngol*; 2005. Poster #898
35. Turner CW, Souza PE, Forget LN. Use of temporal envelope cues in speech recognition by normal and hearing-impaired listeners. *J. Acoust. Soc. Am* 1995;97:2568–2576. [PubMed: 7714274]
36. Van Schijndel NH, Houtgast T, Festen JM. Effects of degradation of intensity, time, or frequency content on speech intelligibility for normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am* 2001;110:529–542. [PubMed: 11508977]

Figure 1.
Illustration of the three spectral shift conditions for Experiment 1.

Figure 2.

Mean vowel recognition performance (across subjects) for experimental conditions in Experiment 1, as a function of test session (day). Baseline and follow-up performance is shown for 8-channel, spectrally-matched (0 mm shift), sine-wave vocoded speech (black triangles). Mean performance is shown on each of the five test days for 8-channel, sine-wave vocoded speech with a 3.6 mm (white circles) or 6 mm shift (white triangles). Data for 8-channel, sine-wave vocoded speech with an 8.3 mm shift (gray squares) are from Fu et al. (2005). The lower dashed line shows chance level performance (8.33%). The error bars show one standard deviation.

Figure 3. Mean percent of F1 and F2 information transmitted (across subjects) on Day 1 and Day 5 for experimental conditions in Experiment 1. The error bars show one standard deviation.

Figure 4.
Illustration of speech processing conditions for Experiment 2.

Figure 5.

Mean performance (across subjects) for the experimental conditions in Experiment 2, as a function of test session (day). The filled symbols show performance for the exclusive exposure group, and the open symbols show performance for the mixed exposure group. The 8-channel, severe shift data for the exclusive exposure group (gray circles) are from Fu et al. (2005). The lower dashed line shows chance level performance (8.33%). The error bars show one standard deviation. Data points have been slightly offset for better visual clarity.

Exclusive exposure

Mixed exposure

Figure 6.

Mean percent of F1 and F2 information transmitted (across subjects) on Day 1 and Day 5 for the exclusive and mixed exposure subject groups in Experiment 2. The error bars show one standard deviation.

Table 1

Analysis band and sine-wave carrier parameters for Experiment 1.

Channel	Analysis bands			Sine-wave carriers								
				3.6 mm shift			6 mm shift			8.3 mm shift		
	Corner Frequency (Hz)	Center Frequency (Hz)	Greenwood Distances from Apex (mm)	Center Frequency (Hz)	Greenwood Distance from Apex (mm)	Amount of shift toward Base (mm)	Center Frequency (Hz)	Greenwood Distance from Apex (mm)	Amount of shift toward Base (mm)	Center Frequency (Hz)	Greenwood Distance from Apex (mm)	Amount of shift toward Base (mm)
1	200–359	272	6.7	544	10.3	3.6	805	12.7	6.0	1168	15	8.3
2	359–591	464	9.4	763	12.3	2.9	1108	14.7	5.3	1585	17	7.6
3	591–930	744	12.2	1052	14.3	2.1	1507	16.7	4.5	2136	19	6.8
4	930–1426	1155	14.9	1433	16.3	1.4	2033	18.7	3.8	2863	21	6.1
5	1426–2149	1754	17.7	1936	18.3	0.6	2726	20.7	3.0	3821	23	5.3
6	2149–3205	2629	20.4	2598	20.3	−0.1	3640	22.7	2.3	5084	25	4.6
7	3205–4748	3905	23.2	3471	22.3	−0.9	4845	24.7	1.5	6748	27	3.8
8	4748–7000	5769	25.9	4622	24.3	−1.6	6433	26.7	0.8	8942	29	3.1

Results from one-way RM ANOVAs (with test session as factor) for speech processing conditions in Experiment 1. Vowel recognition performance was compared between Day 1 and Day 5 for the 0 mm shift condition, and across all five days for the three shift conditions. The percent of vowel feature information transmitted was compared between Day1 and Day 5 for all conditions. Significant differences are shown in bold type.

Table 2

Condition	Percent correct			Percent vowel feature information transmitted					
	df	F	p-value	F1		F2		p-value	
Shift (mm)				df	F	df	F		
0 (Matched)	1,11	6.072	0.031	1,11	0.123	1,11	7.443		0.02
3.6	4,12	27.968	<0.001	1,3	27.990	1,3	17.491		0.025
6	4,12	84.152	<0.001	1,3	79.288	1,3	69.020		0.004
8.3	4,12	1.779	0.198	1,3	1.685	1,3	0.0724		0.805

Table 3

Formant categories for vowel feature analysis used to calculate the percent of information transmitted.

	F1 (Height)	F2 (Place)
HAD	Low	Front
HOD	Low	Back
HAWED	Middle	Back
HEAD	Middle	Front
HAYED	Middle	Front
HEARD	Middle	Central
HID	High	Front
HEED	High	Front
HOED	Middle	Back
HOOD	High	Back
HUD	Middle	Central
WHO'D	High	Back

Table 4

Analysis band and sine-wave carrier parameters for Experiment 2.

Channel	8 channels, moderate shift						8 channels, severe shift					
	Analysis bands			Sine-wave carriers			Analysis bands			Sine-wave carriers		
	Corner Frequency (Hz)	Center Frequency (Hz)	Greenwood Distances from Apex (mm)	Center Frequency (Hz)	Greenwood Distance from Apex (mm)	Amount of shift toward Base (mm)	Corner Frequency (Hz)	Center Frequency (Hz)	Greenwood Distances from Apex (mm)	Center Frequency (Hz)	Greenwood Distance from Apex (mm)	Amount of shift toward Base (mm)
1	455–731	580	10.7	1168	15	4.3	200–359	272	6.7	1168	15	8.3
2	731–1134	913	13.4	1585	17	3.6	359–591	464	9.4	1585	17	7.6
3	1134–1723	1401	16.2	2136	19	2.8	591–930	744	12.2	2136	19	6.8
4	1723–2587	2114	18.9	2863	21	2.1	930–1426	1155	14.9	2863	21	6.1
5	2587–3844	3156	21.7	3821	23	1.3	1426–2149	1754	17.7	3821	23	5.3
6	3844–5680	4675	24.4	5084	25	0.6	2149–3205	2629	20.4	5084	25	4.6
7	5680–8361	6894	27.2	6748	27	–0.2	3205–4748	3905	23.2	6748	27	3.8
8	8361–12275	10133	29.9	8942	29	–0.9	4748–7000	5769	25.9	8942	29	3.1

Channel	16 channels, severe shift					
	Analysis bands			Sine-wave carriers		
	Corner Frequency (Hz)	Center Frequency (Hz)	Greenwood Distance from Apex (mm)	Center Frequency (Hz)	Greenwood Distance from Apex (mm)	Amount of shift toward Base (mm)
1	200–272	234	6.0	1080	14.5	8.5
2	272–359	313	7.4	1262	15.5	8.1
3	359–464	409	8.8	1471	16.5	7.7
4	464–591	524	10.1	1710	17.5	7.4
5	591–744	664	11.5	1985	18.5	7.0
6	744–930	833	12.9	2301	19.5	6.6
7	930–1154	1036	14.2	2663	20.5	6.3
8	1154–1425	1283	15.6	3079	21.5	5.9
9	1425–1752	1580	17.0	3557	22.5	5.5
10	1752–2147	1940	18.3	4106	23.5	5.2

Channel	16 channels, severe shift						
	Analysis bands			Sine-wave carriers			
	Corner Frequency (Hz)	Center Frequency (Hz)	Greenwood Distance from Apex (mm)	Center Frequency (Hz)	Greenwood Distance from Apex (mm)	Amount of shift toward Base (mm)	
11	2147–2625	2375	19.7	4736	24.5	4.8	
12	2625–3202	2900	21.1	5459	25.5	4.4	
13	3202–3900	3534	22.4	6289	26.5	4.1	
14	3900–4742	4301	23.8	7243	27.5	3.7	
15	4742–5761	5228	25.2	8337	28.5	3.2	
16	5761–7000	6352	26.6	9593	29.5	2.9	

Table 5

Timeline for Experiment 2. Note that for the mixed exposure group, subjects were exposed to 3–4 blocks of each speech processing condition presented in random order during the test session.

	Day 1		Day 2	Day 3	Day 4	Day 5
	<i>Baseline: Unprocessed; 8-ch, matched</i>	<i>Adaptation: 8-ch, severe shift 4 blocks</i>	<i>Adaptation: 8-ch, severe shift 4 blocks</i>	<i>Adaptation: 8-ch, severe shift 4 blocks</i>	<i>Adaptation: 8-ch, severe shift 4 blocks</i>	<i>Follow-up: 8-ch, matched, Unprocessed;</i>
Exclusive exposure: 8-ch, severe shift (from Fu et al., 2005)						
Exclusive exposure: 16-ch, severe shift	<i>Baseline: Unprocessed; 8-ch, matched 16-ch, matched</i>	<i>Adaptation: 16-ch, severe shift 4 blocks</i>	<i>Adaptation: 16-ch, severe shift 4 blocks</i>	<i>Adaptation: 16-ch, severe shift 4 blocks</i>	<i>Adaptation: 16-ch, severe shift 4 blocks</i>	<i>Adaptation: 16-ch, severe shift 4 blocks</i>
	<i>Adaptation: 16-ch, severe shift 4 blocks</i>					
Mixed exposure: 8-, 16-ch, severe shift; 8-ch, moderate shift	<i>Baseline: Unprocessed; 8-ch, matched 16-ch, matched</i>	<i>Adaptation: 16-ch, severe shift 3–4 blocks; 8-ch, severe shift 3–4 blocks; 8-ch, moderate shift 3–4 blocks</i>	<i>Adaptation: 16-ch, severe shift 3–4 blocks; 8-ch, severe shift 3–4 blocks; 8-ch, moderate shift 3–4 blocks</i>	<i>Adaptation: 16-ch, severe shift 3–4 blocks; 8-ch, severe shift 3–4 blocks; 8-ch, moderate shift 3–4 blocks</i>	<i>Adaptation: 16-ch, severe shift 3–4 blocks; 8-ch, severe shift 3–4 blocks; 8-ch, moderate shift 3–4 blocks</i>	<i>Adaptation: 16-ch, severe shift 3–4 blocks; 8-ch, severe shift 3–4 blocks; 8-ch, moderate shift 3–4 blocks</i>
	<i>Adaptation: 16-ch, severe shift 3–4 blocks; 8-ch, severe shift 3–4 blocks; 8-ch, moderate shift 3–4 blocks</i>					

Results from one-way RM ANOVAs (with test session as factor) for speech processing conditions in Experiment 2. Vowel recognition performance was compared across all five days for all shift conditions. The percent of vowel feature information transmitted was compared between Day 1 and Day 5 for all shift conditions. Significant differences are shown in bold type.

Table 6

	Condition	Percent correct			Percent vowel feature information transmitted					
		df	F	p-value	F1			F2		
Exclusive exposure	8-ch, severe shift	4,12	1.779	0.198	1,3	1.685	0.285	1,3	0.08724	0.805
	16-ch, severe shift	4,12	0.635	0.647	1,3	0.530	0.519	1,3	9.700	0.053
Mixed exposure	8-ch, moderate shift	4,12	14.976	<0.001	1,3	4.33	0.129	1,3	25.056	0.015
	8-ch, severe shift	4,12	4.654	0.017	1,3	0.343	0.865	1,3	0.0017	0.969
	16-ch, severe shift	4,12	8.078	0.002	1,3	4.786	0.117	1,3	11.783	0.041