J Am Acad Audiol 15:281–299 (2004)

Effect of Frequency Boundary Assignment on Vowel Recognition with the Nucleus 24 ACE Speech Coding Strategy

Marios S. Fourakis*
John W. Hawks†
Laura K. Holden‡
Margaret W. Skinner‡
Timothy A. Holden‡

Abstract

Two speech processor programs (MAPs) differing only in electrode frequency boundary assignments were created for each of eight Nucleus 24 Cochlear Implant recipients. The default MAPs used typical frequency boundaries, and the experimental MAPs reassigned one additional electrode to vowel formant regions. Four objective speech tests and a questionnaire were used to evaluate speech recognition with the two MAPs.

Results for the closed-set vowel test and the formant discrimination test showed small but significant improvement in scores with the experimental MAP. Differences for the Consonant-Vowel Nucleus-Consonant word test and closed-set consonant test were nonsignificant. Feature analysis revealed no significant differences in information transmission. Seven of the eight subjects preferred the experimental MAP, reporting louder, crisper, and clearer sound.

The results suggest that Nucleus 24 recipients should be given an opportunity to compare a MAP that assigns more electrodes in vowel formant regions with the default MAP to determine which provides the most benefit in everyday life.

Key Words: Cochlear implant, frequency boundaries, MAPs, questionnaire, speech tests

Abbreviations: ACE = advanced combination encoder speech coding strategy; CI = cochlear implant; CIS = continuous interleaved sampler speech coding strategy; CF = center frequency; CNC = Consonant-Vowel Nucleus-Consonant; F0 = fundamental frequency; F1, F2, F3 = first, second, third formant; /hVd/ = /h Vowel d/; MAP = speech processor program; pps = pulses per second; rms = root mean square; SPEAK = Spectral Peak speech coding strategy

Sumario

Se crearon dos programas de procesamiento del lenguaje (MAP) que diferían sólo en la asignación de límites de frecuencia en el electrodo, para ocho portadores del implante coclear Nucleus 24. El MAP por defecto utilizó límites de frecuencia típicos, y el MAP experimental reasignó un electrodo adicional a las regiones de formantes vocálicos. Se utilizaron cuatro pruebas de lenguaje objetivas y un cuestionario para evaluar el reconocimiento del lenguaje con los dos MAP.

*Ohio State University; †Kent State University; ‡Washington University School of Medicine in St. Louis

Reprint requests: John W. Hawks, Ph.D., Kent State University, School of Speech Pathology and Audiology, P.O. Box 5190, Kent, Ohio 444242; Phone: 330-672-2672; Fax: 330-672-2643; E-mail: jhawks@kent.edu

Parts of this paper were presented at the American Academy of Audiology Convention on April 20, 2002, in Philadelphia, PA.

This research was supported by Grant R01 DC00581 from the National Institute on Deafness and Other Communication Disorders.

Los resultados de la prueba de lista cerrada de vocales y la prueba de discriminación de formantes mostró una pequeña pero significativa mejoría en los puntajes con el MAP experimental. Las diferencias para la prueba de palabras Consonante-Núcleo Vocálico-Consonante y para la prueba de consonantes de lista cerrada no fueron significativas. El análisis de rasgos no mostró diferencias significativas en la transmisión de información. Siete de los ocho sujetos prefirieron el MAP experimental, reportando sonidos más fuertes, más vivos y más claros.

Los resultados sugieren que los receptores de un Nucleus 24 deberían recibir la oportunidad de comparar su MAP por defecto con un MAP que asigna más electrodos en la región de formantes vocálicos, y así poder determinar cuál aporta el mayor beneficio en su vida diaria.

Palabras Clave: Implante coclear, limites de frecuencia, MAP, cuestionario, Pruebas de Lenguaje

Abreviaturas: ACE = Estrategia de codificación del lenguaje: codificador avanzado de combinación; CI = Implante coclear; CIS = Estrategia de codificación del lenguaje: muestreador continuo de interpolación; CF = Frecuencia central; CNC = Consonante-Núcleo Vocálico-Consonante; F0 = Frecuencia fundamental; F1, F2, F3 = primero, segundo, tercer formante; /hVd/ = /h Vocal d/; MAP = programa de procesamiento del lenguaje; pps = pulsos por segundo; rms = raíz media cuadrada; SPEAK = Estrategia de Codificación del Lenguaje de Pico Espectral

Cochlear implants provide recipients with access to auditory information by subjecting an incoming auditory signal to spectral analysis, determining the amount of energy in each of multiple overlapping frequency bands, and translating the energy distribution found to electrical pulses sent to electrodes in the cochlea. The use of digital filter banks, rather than the older fundamental frequency (F0), first formant (F1), and second formant (F2) feature extraction method, to analyze incoming sound has become the standard filtering approach for speech processing strategies in the Nucleus Cochlear Implant System since the introduction of the Spectra 22 speech processor and the Spectral Peak (SPEAK) speech coding strategy in 1994 (Dowell et al, 1987; Skinner et al, 1991; Skinner et al, 1994). At present, this approach continues to be implemented in the Nucleus 24 Cochlear Implant System, which can be programmed with three different speech coding strategies: SPEAK, advanced combination encoder (ACE; Skinner, Arndt, et al, 2002), and continuous interleaved sampler (CIS; Wilson et al, 1995).

The use of filter banks requires that decisions be made with regard to the assignment of frequency boundaries to each filter and, in effect, to each electrode. The importance of these decisions has been illustrated in the literature. A study by Välimaa et al (2002) found that a group of 19 Finnish-speaking, postlinguistically deafened, multichannel cochlear implant (CI) recipients obtained 80% vowel recognition on an open-set nonsense syllable test by 18 months post– initial activation of the cochlear implant. Most vowel confusions occurred when there were minimal spectral differences between vowels. This led the authors to hypothesize that the vowels most often confused with each other may have spectral energy information from the F1 and F2 or the F2 and third formant (F3) being transmitted to the same channel. If different formants occur within one channel, their center frequencies and relative amplitude levels may be obscured, making it difficult or impossible for CI recipients to perceive the differences between vowels that are acoustically similar. The authors, citing the work of Skinner et al (1995, 1997), suggested that it may be important to consider the vowel confusions of individual CI recipients and assign frequency bands to electrodes that would provide optimal recognition of vowels. In the Skinner et al (1995, 1997) studies, subjects using the SPEAK strategy with the Spectra 22 processor were programmed with two frequency boundary assignments differing mainly in the F1 region. In one condition, four filters were assigned to the frequency region between 150 and 950 Hz, while in the other condition there were six filters assigned to the region between 120 and 1080 Hz. In this study, performance on speech perception tasks was significantly improved with the assignment of more electrodes to the F1 region.

The issue of frequency boundary assignment is further complicated by two additional limitations. First, electrode insertion cannot be exactly aligned with the tonotopic organization of the cochlea. Thus, the pitch ranges intended for perception are different from those actually perceived. Harnsberger et al (2001) have argued that CI recipients adapt to these differences over time and that part of the variability in performance is the result of variability in the capacity to adapt. Fu and Shannon (1999a, 1999b, 1999c) have argued that electrode center frequencies (CF) should correspond as closely as possible to the positionally determined frequencies along the cochlea. Furthermore, they have proposed that, to provide the best possible match between electrode CF and cochlear place, the frequency range to be coded should be limited to 500–5000 Hz. Although this issue is of great importance, it is beyond the scope of the current study.

The second limitation is that there is currently no provision for the programming audiologist to assign specific frequency boundary values to electrodes when creating speech processor programs (i.e., MAPs) for individual CI recipients. The assignment of frequency boundaries for the Nucleus 22 and 24 speech processors is accomplished through frequency tables provided by Cochlear Americas as part of the programming software for creation of MAPs for each CI recipient. There are a finite number of tables supplied, a subset of which is made available for recipients with different numbers of active electrodes (for a description, see Skinner, Arndt, et al, 2002). In addition, there is no provision for independently changing filter bandwidths.

Given the limitations brought about by the need to use tables with predetermined frequency boundaries, optimizing an individual's MAP for speech perception through adjustment of filter center frequencies or bandwidths and their assignment to electrodes is limited. However, the existing tables do offer some flexibility in the number of filters (and electrodes) that can be assigned to different frequency regions of the incoming speech signal. For example, default table 8 (recommended by the manufacturer for mapping 20 electrodes with the ACE speech coding strategy) codes the frequency range from 187 to 7937 Hz. It assigns 6 electrodes to the F1 region (187–937 Hz), 7 electrodes to the F2 region (937–2687 Hz), and the remaining electrodes to frequencies above 3000 Hz, with the most basal electrode being assigned to a frequency band from 6937 to 7937 Hz. Given that the SPrint™ processor uses an anti-aliasing low-pass filter during analog-to-digital conversion with a cutoff frequency of 7000 Hz and a slope of 36 dB/octave (van den Honert and Miller, personal communication), information conveyed by the most basal electrode is limited. Between 7000 and 7937 Hz, the filter attenuation is ~6.5 dB. In addition, there is a 7 dB reduction in the microphone response between 6 and 8 kHz. Given these two sources of attenuation, the incoming signal amplitude at 7937 Hz is approximately 13.5 dB lower than at 6000 Hz. This decrease in level above 6000 Hz is coupled with the low amplitude of speech sounds in the region from 6000 to 7937 Hz. Finally, the speech contrasts conveyed by spectral differences in the frequency region from 5000 to 6500 Hz are mainly voiceless fricatives (/f/, /θ/, /s/) and low intensity voiced fricatives (/v/, /ð/, and /z/). Thus, a case can be made that the assignment of the most basal electrodes to frequency bands between 6000 and 8000 Hz may not contribute substantially to speech perception by CI recipients. As an alternative, the use of nondefault tables with a lower, high-frequency limit can result in the assignment of more electrodes to frequencies below 6000 Hz. It is hypothesized that providing additional information from critical formant frequency regions will improve speech perception performance as well as user satisfaction with their cochlear implants.

The objective of the present study was to determine whether vowel recognition can be improved and recognition of frication in consonants retained by reallocating the frequency bands assigned to electrodes within the ACE strategy, given the frequency boundary tables available in the clinical programming software. To evaluate this objective, two MAPs were created for each subject, one using a typical default frequency boundary table that assigns frequencies between 187 and 7937 Hz, and a second, using a different table that limits the highest frequency represented to ~6000 Hz.

METHOD

Subjects

Eight postlinguistically deaf adults who were recipients of the Nucleus 24 Cochlear Implant System and had used the ACE speech coding strategy for at least six months participated in the study. Biographical information is given in Table 1. All subjects were native speakers of American English and were fluent auditory/oral communicators. Average age at implantation was 51.9 years. Four subjects were implanted with a Nucleus 24 contour electrode array (Parkinson et al, 2002) with insertion to the intended depth as indicated by the silastic marker. The other four subjects were implanted with a Nucleus 24 straight electrode array. The surgeon reported that not all of the 32 bands in the electrode array (i.e., 22 active electrodes plus 10 supporting rings) were inside the cochlea for these four subjects; that is, the array was not implanted to its intended insertion depth. Subject 8 (S8), however, was the only subject who did not have all 22 active electrodes inserted into the cochlea. For S8, 20 active electrodes were inserted.

Cochlear Implant System

The Nucleus 24 Cochlear Implant System used for this study included an internal device (CI24M & CI24R), a body-worn speech processor (SPrint™), and a Windows programming and diagnostic system (WinDPS, version 116). The internal device consisted of a receiver/stimulator connected to a straight or contour array of 22 intracochlear electrodes and two extracochlear electrodes (a plate electrode on the implant package and a ball electrode on a lead that is placed under the temporalis muscle during surgery). A brief description of the Nucleus 24 implant is given by Vandali et al (2000). For this study, the intracochlear electrodes were stimulated in a monopolar configuration, and the extracochlear electrodes were connected together as the ground electrode.

Speech Processor Programs (MAPs)

All subjects used 25 μs/phase, monopolar stimulation with the ACE strategy. Prior to the study, thresholds and growth of loudness to maximum acceptable loudness levels were obtained on each electrode (as described by Sun et al, 1998) to determine whether clinically significant changes had occurred since the last speech processor MAP was created. If changes had occurred, small adjustments in MAP minimum and maximum stimulation levels were made.

Other MAP parameters are provided in Table 2. Of these parameters, the choice of base level affects the size of the input dynamic range (i.e., 30 dB and 41.5 dB for base levels of 4 and 1, respectively). The Q-value is the function for mapping of coded input sound amplitude onto an individual's electrical dynamic range for each electrode (for a description of base level and Q-value, see Skinner et al, 1999). The number of channels

Table 1. Biographical Information as Well as Type of Electrode Array and Its Depth of Insertion

Subject	Sex	Etiology	Duration of Deafness (years)	Age at Implantation	Type of Electrode Array	Depth of Insertion (# bands inside)	Length of Implant Use (months)
1	F	Unknown	15	37	Straight	30	15
2	M	Meningitis	5	34	Contour	Full	12
3	M	Meniere's/ Noise Exposure	2	71	Contour	Full	12
4	F	Genetic	7	48	Contour	Full	11
5	M	Meningitis	2	36	Straight	24	30
6	M	Unknown	2	64	Straight	28	36
7	F	Unknown	0.5	65	Contour	Full	7
8	F	Otosclerosis	10	60	Straight	20	25

Table 2. SPrint™ Speech Processor MAP Parameters Used by Each Subject for the Default and Experimental Frequency Table Assignments

Subject	Base Level	Q-value	# channels	# active electrodes	# maxima	Rate/channel (pps)	Total rate/cycle (pps)
1	1	20	20	20	12	1200	14,200
2	4	25	19	19	8	1800	14,400
3	4	20	19	17	8	1800	14,400
4	4	20	17	16	12	900	10,800
5	4	25	20	20	8	1800	14,400
6	4	20	20	20	8	1800	14,400
7	4	20	20	20	8	900	7200
8	4	20	16	16	8	1800	14,400

is the number of frequency bands into which the incoming sound is filtered. The number of channels equals the number of active electrodes chosen for stimulation except when more than one channel is assigned to an electrode. For six of the eight subjects, each channel was assigned to a different electrode. For the other two subjects, one channel per electrode was assigned, except for one or two electrodes for which two channels were assigned to one electrode. For S3, channels 21 and 20 were assigned to electrode 21, and channels 17 and 16 were assigned to electrode 16. This double channel mapping was done to eliminate facial nerve stimulation and maintain a wide frequency range. For S4, channels 11 and 10 were assigned to electrode 10 due to pain sensation when electrode 11 was stimulated. Both subjects had used these MAPs since their initial stimulation with the implant (approximately one year prior to the start of this study). This assignment of two channels to one electrode results in stimulation of that electrode in response to a wider frequency band of incoming sound than when one channel is assigned to one electrode. If either or both of the two channels are chosen as a maxima during sound processing, that electrode is stimulated.

Maxima are the channels with the highest amplitudes of incoming signal, and their corresponding electrodes are stimulated during each cycle with the ACE strategy (Vandali et al, 2000). As shown in Table 2, subjects used either 8 or 12 maxima. Stimulation rates on each channel ranged between 900 and 1800 pulses per second (pps) per channel, with a total stimulation rate between 7200 and 14,400 pps/cycle. These rates were selected during the clinical fitting process (S1–4, S7) or during a previous study of the effect of rate on speech recognition (S5, S6, S8; Holden et al, 2002).

Subjects 3, 4, 6, 7, and 8 used the default assignment of frequency boundaries to channels prior to the study. Subjects 1, 2, and 5 participated in a pilot study comparing the default and experimental assignment of frequency boundaries. These three subjects chose to use the experimental assignment, which they had used for approximately six months prior to this study, because they reported that speech and other sounds were clearer with it than with the default assignment.

For the present experiment, two MAPs were created for each subject using the clinical programming software (version R116) and the SPrint™ speech processor. This processor uses FFT (Fast Fourier Transform) based filterbanks that are typically 180 Hz wide with FFT bin spacing of typically 125 Hz. For this reason, the bandwidths assigned to channels are multiples of this spacing. The MAPs were identical except for the frequency tables used to assign frequency bands to electrodes. One MAP was created using the software's default assignment of frequency bands to channels (hereafter termed the "default MAP"), and the other MAP was created using a different assignment of frequency bands (hereafter termed the "experimental MAP"). The total range of frequencies in the default MAP was always 187 to 7937 Hz, regardless of the number of channels of stimulation. For the experimental MAP, 187 Hz was always the lowest frequency, while the high-

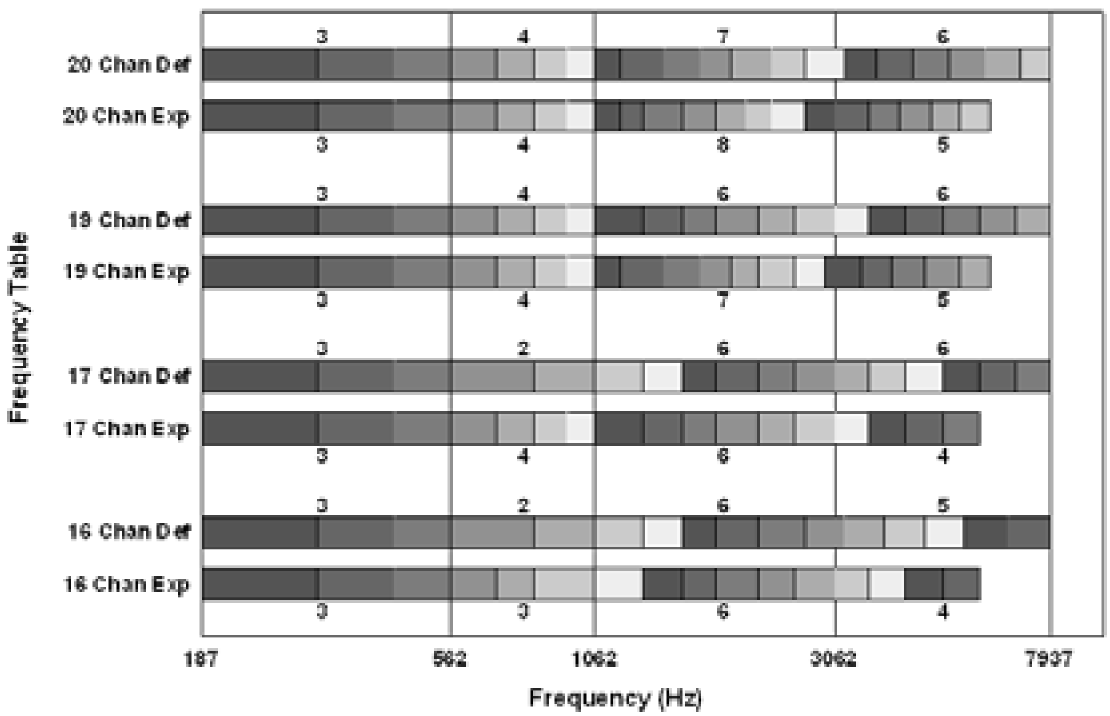

Figure 1. Frequency assignments for the default and experimental MAPs for 20, 19, 17, and 16 channels. Each shaded area represents one channel. The width of the shaded area represents the frequency band for that channel.

est frequency was either 6062 Hz (for MAPs with 20 or 19 channels) or 5812 Hz (for MAPs with 17 or 16 channels). Figure 1 shows how the frequency bands varied for the default and experimental MAPs with 20, 19, 17, and 16 channels. Note that the number of channels and their bandwidths for frequencies between 187 Hz and 562 Hz are identical across all MAPs and electrode configurations. For 20 and 19 channels, frequency bands remain identical for the default and experimental MAPs between 187 and 1062 Hz, and for 20 channels, the frequency band between 1062 and 1187 Hz is also identical. For experimental MAPs, there was a shift to a more basal electrode compared to the default MAPs for frequency bands higher than 1187 Hz (20 channels), 1062 Hz (19 channels), and 562 Hz (17 and 16 channels). This shift results in more electrodes assigned to the F2/F3 regions for 19- and 20-channel experimental MAPs, and to the F1 region for 16- and 17-channel experimental MAPs. In addition, note that a number of bands in the experimental MAP had narrower frequency ranges than in the default MAP.

Warble-Tone Stimuli for Sound-Field Threshold Levels

Warble tones were used to determine the sound-field threshold levels for MAPs with the default and experimental frequency tables. The warble tones (centered at 250, 500, 750, 1000, 1500, 2000, 3000, 4000, and 6000 Hz) were sinusoidal carriers modulated with a triangular function over the standard bandwidths recommended for use in the sound field by Walker et al (1984). The modulation rate was 10 Hz. A conversion from dB SPL to dB HL in the sound field was made using data obtained by Pascoe (1975) and Skinner (1988) at Central Institute for the Deaf (CID). For this conversion, the following values are subtracted from dB SPL values: 22, 11, 9, 7, 5, 3, -2, -3, and 6 dB at 250, 500, 750, 1000, 1500, 2000, 3000, 4000, and 6000 Hz, respectively.

Speech Tests

Four tests were chosen for this evaluation: closed-set vowel and consonant tests, an open-set word test, and a synthetic vowel test. All speech tests, except the synthetic vowel test, had been used in previous research studies with adult cochlear implant recipients. The closed-set vowel test included a set of 11 vowels recorded in an /h Vowel d/ context (/a, æ, ɛ, ɚ, eɪ, ɪ, i, oʊ, ʊ, ʌ, u/). One example of each vowel from two male and two female speakers of American English were selected from a recording made by Hillenbrand et al (1995). For the male and female speakers separately, two examples of each token were selected on the basis of their similarity in fundamental frequency. Each vowel token was digitized to a computer audio file, and all tokens were normalized for amplitude using their rms (root mean square) levels.

The closed-set consonant test included 16 consonant stimuli in an /a Consonant a/ context: /p, t, k, b, d, g, f, s, ʃ, m, n, v, z, ʒ, r, l/. These consonants were part of the original set of 19 consonants recorded and described by Van Tasell et al (1992). One example of each consonant was recorded by three male and three female speakers. For the present study, one example of each consonant spoken by two male and two female speakers was selected. Only two of the female voices were used in the test materials because the third had a much different fundamental frequency. For the male speakers, most of the examples were from speakers four and six; only a few tokens from speaker five were included because of his breathy voice. These choices were made

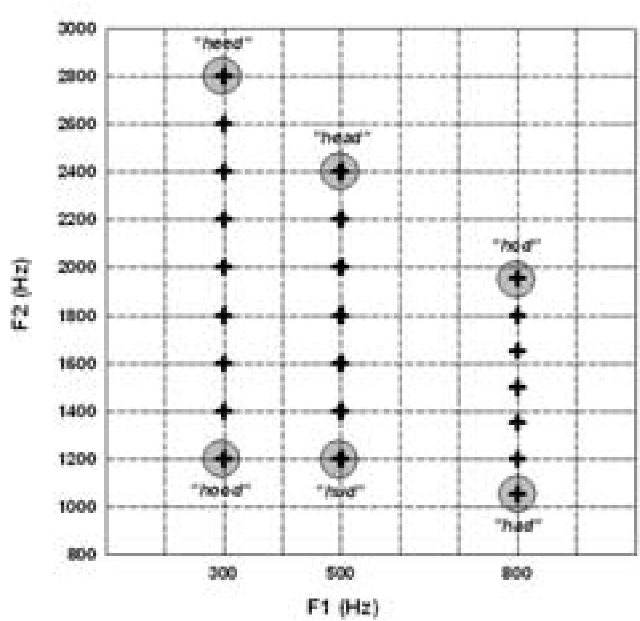

Figure 2. The F1 by F2 center frequencies of the four-formant vowel tokens used in the synthetic vowel discrimination test.

so that the two tokens of each consonant spoken by the males sounded perceptually similar.

The third test was the Consonant-Vowel Nucleus-Consonant (CNC) word test (Peterson and Lehiste, 1962) spoken by a male talker and recorded for Cochlear Americas on compact disc. This recording is part of the Minimum Speech Test Battery for Adult Cochlear Implant Users and is accepted as part of the minimal auditory battery for cochlear implant research (Luxford and Ad Hoc Subcommittee, 2001). The fourth test was a synthetic vowel formant discrimination test. For this test, three sets of synthetic four- formant vowel tokens were constructed with a fundamental frequency contour typical for a male talker. To evaluate discrimination ability over a wide range of vowel qualities, each set consisted of tokens using a single F1 center frequency and a range of appropriate F2 center frequencies (Set 1: F1 = 300 Hz, F2 = 1200–2800 Hz in 200 Hz steps; Set 2: F1 = 500 Hz, F2 = 1200–2400 Hz in 200 Hz steps; and Set 3: F1 = 800 Hz, F2 = 1050–1950 Hz in 150 Hz steps). This design yielded a total of 23 tokens. As shown in Figure 2, the vowels in "hood" and "heed" bound the F2 range when F1 = 300 Hz; the vowels in "hud" and "head" bound the F2 range when F1 = 500 Hz; and the vowels in "had" and "hod" bound the F2 range when F1 = 800 Hz. The F3 center frequencies were computed on the basis of the F1 and F2 frequencies using the algorithm described in Nearey 1989. The F4 center frequency was fixed at 3700 Hz for all tokens. All formant bandwidths were computed using the algorithm described in Hawks and Miller (1995). The synthesis employed the cascade branch of the KLSYN88A software synthesizer (Klatt and Klatt, 1990) using a 10 kHz sampling rate with 12 bit precision. All tokens were 250 msec long with 20 msec amplitude ramps. Overall amplitude was normalized across tokens within ±1 dB. The fundamental frequency contour used an initial value of 145 Hz, linearly interpolated to 105 Hz at the end of the stimulus.

Equipment/Test Environment

All speech test materials were presented to subjects in a double-walled sound booth (IAC; model 1204-A; 254 cm x 264 cm x 198 cm) through a loudspeaker placed at ear-level height, at 0 degrees azimuth, and at a distance of 1.5 meters from the center of subjects' heads in their absence. The test materials were presented via an IBM compatible, Pentium II computer, interfaced to a CD-laser disc player (Pioneer, model CLD-V 2600), which controlled a mixing and attenuation network (Tucker-Davis Technologies) to present sound through a power amplifier (Crown model D-150) and loudspeaker (Urei; model 809). All speech tests were recorded on compact disc or stored as wave files on the hard disk. The sound pressure level (SPL) of the stimuli was measured with the microphone (Bruel & Kjaer, Model 4155) of the sound level meter (Bruel & Kjaer, Model 2230) at the position of the subject's head during testing. The overall SPL of the words, consonants, and vowels was taken as the average of peaks on the slow, rms, linear scale. Each subject's speech processor sensitivity and volume control settings were identical across all test sessions. These settings were those that each subject used most often in everyday life.

Questionnaire

The subjects responded to a questionnaire about their use of the default and experimental MAPs in everyday life. Each chose the MAP that provided them with the best understanding of speech for 17 listening situations, as well as the best quality for music and detection of environmental sounds (see Table 3). They could also respond that there was no difference between the two MAPs. Subjects were then asked to choose which MAP they preferred to use in everyday life and to comment on subjective differences between the two MAPS.

Research Design and Procedures

A four-phase test design was used with each phase lasting three weeks. For the first phase, half the subjects (S4, S5, S7, S8) used the default MAP and half (S1, S2, S3, S6) used the experimental MAP. At the end of each phase, subjects' speech processors were pro-

Table 3. Listening Situations Evaluated on the Questionnaire

1. At home, when someone is speaking to me from about two to three feet away.
2. At home, when someone is speaking to me from across the room.
3. At home, when someone is speaking to me from another room.
4. At a restaurant, when sitting across from my spouse or friend.
5. Several friends or family members talking around the dinner table.
6. Talking with a friend or spouse outdoors.
7. Listening to a speaker at church or in a meeting (without an FM system).
8. Conversations with coworkers and supervisor.
9. Talking with a particularly soft-spoken person.
10. A waitress at a restaurant.
11. A cashier at the grocery store or department store.
12. A child (6–10 years old).
13. Understanding over the telephone.
14. Listening to your phone message machine.
15. Listening to the television.
16. Listening to news on the radio.
17. Understanding the lyrics of music.
18. Quality of music.
19. Detection of soft environmental sounds (i.e., microwave beeping, computer running, door bell, knocking at door).

grammed with the alternate MAP so that they listened with each of the MAPs during two phases. To familiarize subjects with the speech tests, practice lists were administered one week after the first MAP was placed on the processor and before the first test session. Sound-field thresholds were obtained during the first test session of each phase. Speech testing was done during the last two sessions of each phase. At the end of the study, both MAPs were placed in two of the four program memory locations in the SPrint™ processor so that subjects could compare speech and other sounds in everyday life, especially the situations listed in Table 3. To eliminate potential bias, subjects were not told which MAP was in location one and which was in location two. Subjects had both MAPs on their processors for a two-week period. During this time, they made the comparisons needed for responding to the questionnaire.

The level at which the speech tests were presented was lower than the typical 70 dB SPL presentation level used to test CI recipients (Luxford and Ad Hoc Subcommittee, 2001). At Washington University in St. Louis, CI recipients' speech processors are programmed so that soft and even very soft speech is audible. For this study, group mean warble-tone sound-field threshold levels were below 21 dB HL for each measured frequency (see Table 5). Because subjects could detect very soft sounds and to prevent the highest performers from obtaining scores over 90% correct, speech tests were presented at relatively low levels.

At each test session, one ordering of the female vowel tokens and consonant tokens and an ordering of the male vowel tokens and consonant tokens were presented at 45 dB SPL. For the vowel and consonant tests, the order of presentation of the female and male lists was counterbalanced across sessions and subjects. At each test session, two lists of CNC words were presented, one each at 50 and 60 dB SPL. The order of presentation levels was also counterbalanced across sessions and subjects. One synthetic vowel test was also presented at 50 dB SPL during each test session.

For the synthetic vowel test, each token within a given F1 set was paired with each of the other tokens within the same set, yielding a total of 78 stimulus pairs. All pairs were presented in both AB and BA orders in each listening session, yielding a minimum of 156 stimulus pair presentations per session. This is considered the minimum presentation number because the computer compared subjects' responses to the two orderings of each stimulus pair. If the responses were different (i.e., one response was "same" and one response was "different"), a third response was elicited by means of an additional presentation to determine the session response for that pair. A different randomized order was used for each session. A simple same/different task was used with the subject touching the appropriate box on the touch screen for each trial. Subjects were free to hear repetitions of stimulus pairs at will. There were no time constraints placed on the response interval, with the next trial initiated approximately one second after a response was made. Subjects listened to all stimulus pairs at each of the eight test sessions (four times with the default MAP and four times with the experimental MAP).

In summary, subjects responded with each MAP to a total of four lists of closed-set vowels spoken by each of two talkers, four lists of closed-set consonants spoken by each of two talkers, four lists of CNC words presented at each of the two levels, and four synthetic vowel tests.

Data Analysis

Analyses of variance (ANOVAs) were conducted for each of the speech tests to determine if there was a significant difference in mean score between the default and experimental MAPs for the group. If there was a significant effect of MAP and a significant Subject x MAP interaction, then follow-up comparisons were conducted for each subject.

ANOVA was also used to determine whether there was a significant difference between warble-tone thresholds with the default and experimental MAPs. In these analyses, all effects, including subjects, were considered fixed, and the highest-order interaction was assumed to be zero.

For each MAP (default and experimental), group and individual-subject scores in percent correct were calculated for the closed-set vowel and consonant tests and the open-set CNC word test. For the CNC words at the two presentation levels and using default as well as experimental MAPs, percent correct scores were generated from subjects' responses for the initial consonants, medial vowels, and final consonants. ANOVA was used in the manner described above to determine whether there was a significant difference in group mean scores for the CNC consonants and vowels as a function of MAP and level.

Subjects' responses for the closed-set and open-set vowels and consonants were transcribed and submitted to information transmission analysis (Miller and Nicely, 1955; Wang and Bilger, 1973). Three vowel features were chosen for examination: (1) the first formant feature, coded into high, mid, and low categories of tongue position during vowel articulation (i.e., 1 = 300–450 Hz; 2 = 451–580 Hz; 3 = 581 Hz or higher); (2) the second formant feature, coding vowels into front and back categories (i.e., 1 = 900–1400 Hz; 2 = 1900 Hz or higher); and (3) a feature specifying whether the second formant changed during the production of the vowel (i.e., 1 = no change [monophthongal vowels]; 2 = change toward higher frequencies [diphthongs like /aɪ/]; and 3 = change toward lower frequencies [diphthongs like /au/]). See Table 4 for the vowel coding system. One consonant feature, frication, was chosen for examination to determine whether elimination of the frequencies above ~6000 Hz were associated with a decrease in information transmitted. The following consonants were coded as fricatives: /ʍ, f, v, θ, ð, ʃ, s, ʒ, z, h, tʃ, ʤ /; all other consonants were coded as nonfricatives.

The percentage of information transmitted for the three vowel features was also calculated for the group for the vowel test, as well as for the vowels in CNC words presented at 50 and 60 dB SPL. It was also calculated for individuals and the group for the consonant feature, frication, in the consonant test, as well as initial and final consonants in CNC words, for both presentation levels. A paired t-test was used to evaluate whether there was a significant difference in information transmitted for each test and feature between the default and experimental MAPs for the group.

Subjects' performance for discrimination of formant changes in the synthetic vowel test was quantified as follows. Session responses to all 78 stimulus pairs were collected from all subjects four times with each of the two MAPs. Each stimulus pair consisted of two tokens differing in F2 frequency. Each response of "different" was scored as one, and "same" as zero. Each score was then weighted by multiplying it by the inverse of the difference in frequency between the F2 values for that stimulus pair. Thus, a "different" response for a stimulus pair with a F2 difference of 200 Hz would have a higher weighted value than for a stimulus pair with an F2 difference of 600 Hz. In this way, the ability to hear smaller differences in F2 change was awarded a proportionately greater score than for hearing only larger differences. This reflects how well F2 differences were discriminated regardless of F1 or the actual F2 values involved. The scores were then multiplied by 100 for ease of data management.

A repeated-measures ANOVA was computed to determine if there was a significant difference in the weighted scores between the default and experimental MAPs for the group and for individuals. The group analysis was calculated using the individual weighted trial scores, such that there were three within-subjects factors arranged in factorial fashion: Subject (8 levels); Condition (2 levels, experimental and default); and Phase (4 levels). Individual analyses were similar to this design, but without the Subject factor.

Table 4. Partial Vowel Feature Matrix Used for Information Transmission Analysis

	Feature		
Vowel	F1	F2	F2 Movement
u	1	1	1
ʊ	1	1	1
a	3	1	1
or	2	1	2
ar	3	1	2
ʌ	3	1	1
ɝ	2	1	1
æ	3	2	1
ɛ	2	2	1
i	1	2	1
ɪ	1	2	1
oʊ	2	1	3
eɪ	2	2	2
aɪ	3	1	2
aʊ	3	1	3
ɔɪ	2	1	2

Note: Some vowels listed have identical feature specifications. The complete matrix would include other features, e.g., duration, which would disambiguate all vowels. See text for more information. F1: first formant, F2: second formant.

RESULTS

Questionnaire

Figure 3 shows a summary of individual and group responses to the questionnaire. Each bar represents the percentage of listening situations for which subjects indicated a preference for the default MAP, the experimental MAP, or no perceived difference between MAPs. Five of the eight subjects (S1, S2, S4, S5, and S6) preferred the experimental MAP in at least 50% of the listening situations on the questionnaire (see Table 3). Three of these five subjects preferred the default MAP in equal to or less than 15% of the listening situations. Two subjects (S7 and S8) preferred the experimental MAP in a few more listening situations than the default MAP, but for approximately 75% of the listening situations, they perceived no difference between the MAPs. Subject 3 perceived no difference between the MAPs in 100% of the listening situations. For the group, the experimental MAP was preferred in 48% of the listening situations, the default MAP was preferred in 9% of the listening situations, and no difference was perceived between the two MAPs in 43% of the listening situations.

The "P" at the top of certain bars in Figure 3 indicates the MAP each subject preferred for use in everyday life. Seven subjects preferred the experimental MAP and continued to use it at the end of the study. Subject 3 could not tell a difference between the two MAPs and, therefore, had no preference. All subjects were asked to give reasons why they preferred the MAP chosen for use in everyday life. Two or more subjects used one of the following terms or phrases to describe their preference for the experimental MAP: "clearer," "crisper," "picks up more sound," "better for understanding speech from a distance," "better for listening to music," "lyrics of music are easier to understand," "better for understanding over the telephone and cellular phone." It is interesting to note that two subjects (S1 and S2) commented that they preferred the default MAP for noisy situations because it "damps" the background sound or "picks up less sound" than the experimental MAP. Six of the eight subjects (S1, S2, S4, S5, S6, and S8) commented that the experimental MAP was louder than the default MAP.

Figure 4 shows the number of subjects in the group who preferred the default MAP or experimental MAP or who perceived no differences between the MAPs for each of the 19 listening situations listed in Table 3. In 18 of the 19 listening situations, the experimental MAP was preferred by more subjects than the default MAP. For 10 of the 19 listening situations, at least half of the subjects preferred the experimental MAP over the default MAP. These were relatively difficult listening situations where no speech reading cues were available (e.g., someone speaking from another room, understanding over the tele-

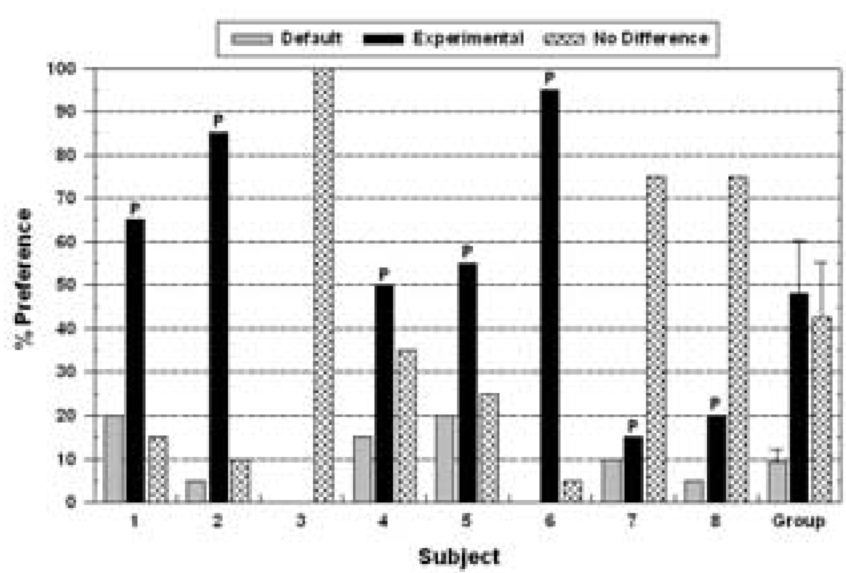

Figure 3. Bars represent the percentage of listening situations listed on the questionnaire for which each subject and the group indicated a preference for either the default MAP (gray bars) or experimental MAP (black bars), or perceived no difference (cross-hatched bars) between the MAPs. The "P" at the top of certain bars denotes the MAP that each subject preferred to use in everyday life at the end of the study.

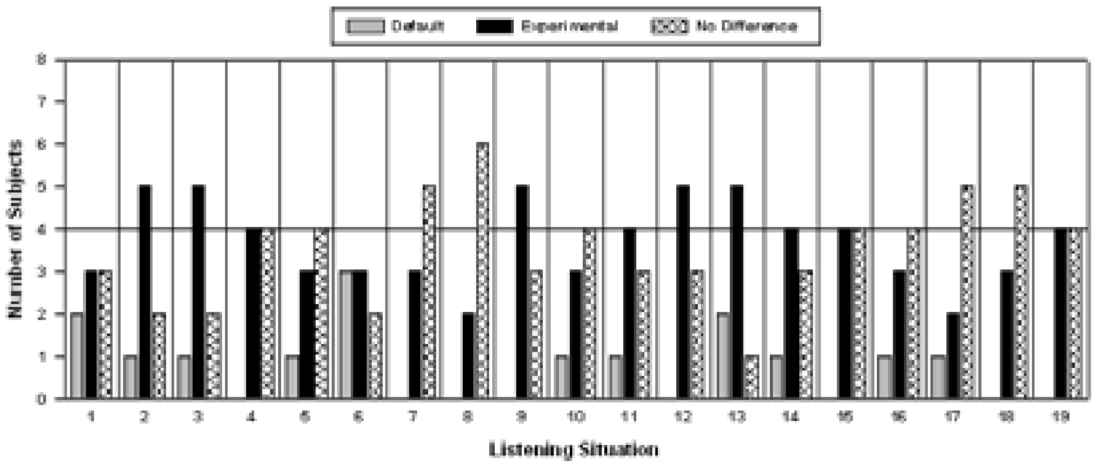

Figure 4. Each bar represents the number of subjects who preferred either the default MAP (gray bars) or experimental MAP (black bars), or who perceived no difference (cross-hatched bars) between the MAPs for the 19 listening situations listed on the questionnaire.

phone, or understanding an answering machine message) or in noise (e.g., waitress at a restaurant or cashier at a grocery store). The majority of subjects also preferred the experimental MAP for listening situations that required increased loudness to understand speech (e.g., listening from across a room or from another room or listening to a soft-spoken person or a child). This coincides with subjects' comments that the experimental MAP was louder than the default MAP. There were four listening situations in which the majority of subjects perceived no difference between the two MAPs. Two of these listening situations pertained to music. Even though the majority of subjects did not perceive a difference in the two MAPs when listening to music, S1 and S4 felt strongly that the experimental MAP provided them with greater music enjoyment than the default MAP.

Warble-Tone, Sound-Field Threshold Levels

Group mean and standard error of the mean for warble-tone, sound-field threshold levels at each frequency are shown for the default and experimental MAPs in Table 5. For both MAPs, group mean thresholds were less than 20 dB HL except at 500 Hz for the default MAP and at 3000 Hz for both MAPs. These thresholds are consistent with an Articulation Index value of close to one. That is, almost all the speech cues in everyday life are audible with both MAPs (e.g., French and Steinberg, 1947; Pavlovic et al, 1985; Pavlovic 1986; Mueller and Killion, 1990). Thresholds at 6000 Hz are significantly higher ($F [1,16] = 16.94$; $p < .001$) with the experimental than the default MAP. This difference is expected because the upper frequency bound with the experimental MAP was limited to 6062 or 5812 Hz. Thus, the amount of energy processed from the 6000 Hz warble tone was less than with the default MAP, which had a substantially higher upper frequency limit. However, the 6000 Hz group mean threshold for the experimental MAP was 17 dB HL, a level that should not compromise audibility of speech sound energy at 6000 Hz in everyday life.

Speech Tests

Group mean percent correct scores for the closed-set vowel test, closed-set consonant test, and CNC word test (scored for both words and phonemes) at the two presentation levels are shown in Figure 5 for both the default and experimental MAPs. The group mean score on the closed-set vowel test was significantly higher with the experimental MAP than with the default MAP ($F [1,32] = 6.27$; $p = .014$). Figure 6 shows individual data for this vowel test for both MAPs. Two subjects (S5 and S8) had significantly higher scores with the experimental MAP than with the default MAP (S5: $F [1,4] = 8.39$; $p = .013$; S8: $F [1,4] = 15.99$; $p = .002$). There were no

Table 5. Group Mean (dB HL) and Standard Error of the Mean for Warble-Tone Thresholds at Each Frequency for the Default and Experimental MAPs

Frequency Table		Warble Tone (Hz) 250	500	750	1000	1500	2000	3000	4000	6000
Default	Mean	16.1	20.9	15.8	18.6	16.9	17.3	21.4	17.4	15.6*
	SEM	1.29	1.67	1.26	1.48	0.99	1.31	0.78	2.01	1.62
Experimental	Mean	16.4	19.4	15.6	18.4	16.5	17.9	20.8	16.9	17.3
	SEM	1.22	1.84	1.67	1.35	0.91	1.09	1.37	3.37	2.21

*Significantly lower threshold for the default MAP at 6000 Hz.

significant male/female talker effects for vowels. Despite the low presentation level (45 dB SPL), the group mean scores with the two MAPs were within a 5% range around 60% correct (chance score: 9.0%).

The group mean score on the closed-set consonant test with the experimental MAP (32.9%) was not significantly higher than with the default MAP (33%). In view of the relatively low group mean scores, the low presentation level of 40 dB SPL limited the

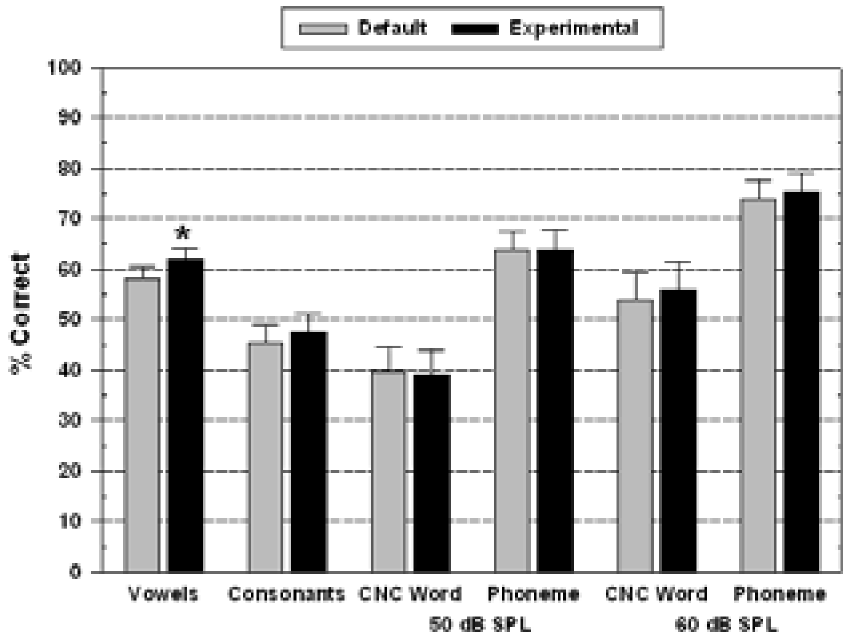

Figure 5. Group mean scores in percent correct for the closed-set vowel and consonants tests and the CNC word test (scored for both words and phonemes at the two presentation levels) for the default and experimental MAPs. Error bars represent one standard error of the mean. Level of significance: * = p ≤ .05.

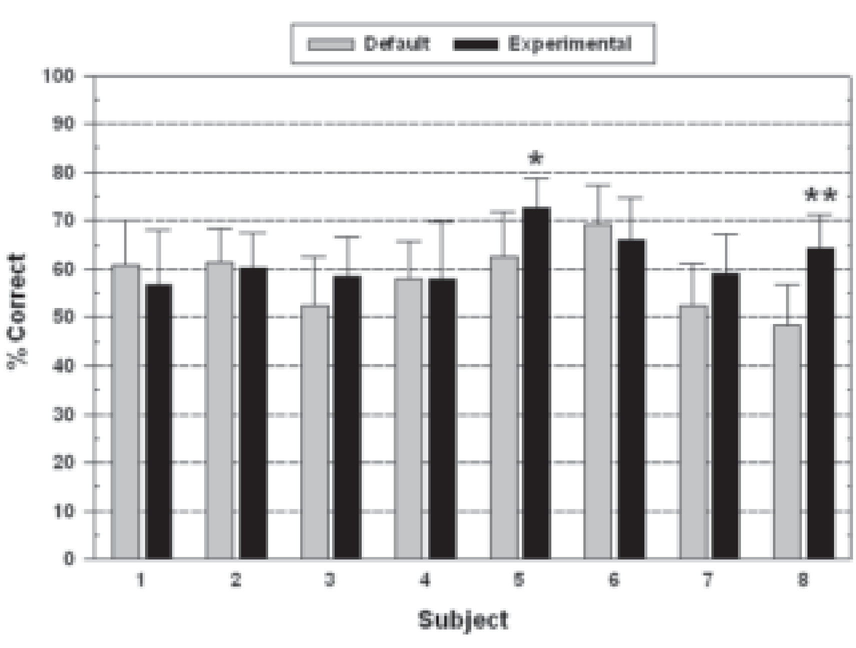

Figure 6. Individual subject's mean scores in percent correct for the closed-set vowel test for both the default and experimental MAPs. Error bars represent one standard deviation. Level of significance: * = p ≤ .05; ** = ≤ .01.

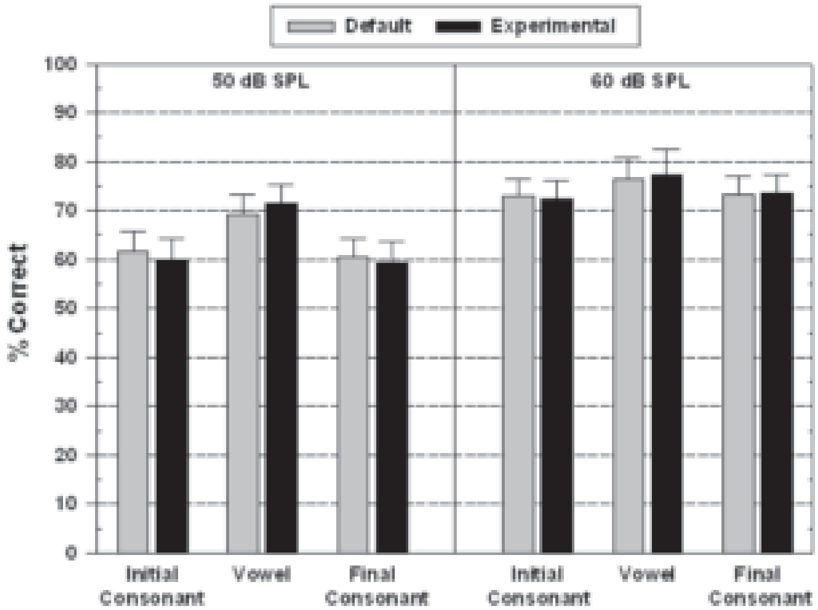

Figure 7. Group mean scores (percent correct) for CNC initial consonants, medial vowels, and final consonants presented at 50 and 60 dB SPL for both the default and experimental MAPs. Error bars represent one standard error of the mean. None of the comparisons were significantly different.

audibility of some consonants (i.e., voiceless stop, /p/; fricatives, /f/, /v/, /s/, /z/; and the nasal, /m/) with both MAPs. It is assumed that audibility was limited because, for both the default and experimental MAPs, the group mean scores were low (i.e., between 9 and 45%). In contrast, the group mean scores for the fricatives, /ʃ/ and /ʒ/ (which are more intense), were between 67 and 94% with the default map (a result consistent with these consonants being audible) and were slightly less with the experimental MAP (by 9 and 4 percentage points for /ʃ/ and /ʒ/, respectively).

The group mean scores for CNC words presented at 50 and 60 dB SPL were analyzed according to whole words as well as phonemes for both MAPs. These scores with standard error of the mean bars are shown in Figure 5. Figure 7 shows the group mean scores and standard error of the mean bars for CNC words analyzed by initial consonants, medial vowels, and final consonants. There was no significant difference in scores between the two MAPs for any of these conditions; however, all scores were significantly higher at 60 dB SPL than 50 dB SPL ($p < .001$).

Analysis of percentage of information transmitted showed the following. For vowels in the closed-set vowel and CNC word tests, there were no significant differences between the two MAPs in percentage of information transmitted for the F1, F2, and F2 movement features. For fricatives in the closed-set consonant and CNC word tests, there were no significant differences between the two MAPs in the percentage of information transmitted; furthermore, for the fricatives /f/, /v/, /s/, the score was either the same or higher with the experimental compared to the default MAP. These results suggest that limiting the frequency range to 6000 Hz was not associated with a consistent, significant decrease in the transmission of frication information for the speech stimuli at the presentation levels used in this study.

Synthetic Vowels

Statistical analysis of the group data for discrimination of the synthetic vowels indicated significantly better performance with the experimental MAP than with the default MAP ($F [1,77] = 7.90$; $p = .006$). The other two main factors in the analysis, Subject ($F [7,71] = 6.47$; $p < .001$), and Phase ($F [3,75] = 6.74$; $p < .001$), also revealed significant differences. The significant Phase factor was apparently due to a gradual performance increase over time, suggesting a slight training effect.

Figure 8 summarizes each subject's mean weighted performance and the group means for the synthetic vowel test. Five of the eight subjects had more accurate vowel discrimination when using the experimental MAP than the default MAP; however, only two subjects had significantly higher scores with the experimental MAP: S1 ($F [1,77] = 6.72$; $p = .011$) and S5 ($F [1,77] = 4.56$; $p = .036$). Only S2 showed a large, significant performance advantage ($F [1,77] = 6.96$; $p = .010$) with the default MAP. These individual differences also brought about a significant

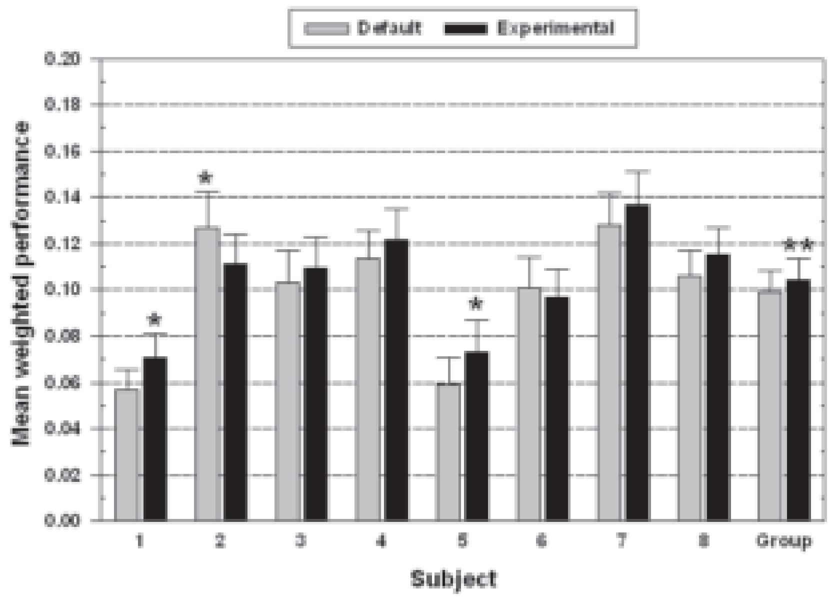

Figure 8. Individual and group mean weighted performance for the synthetic vowel test for both the default and experimental MAPs. Error bars represent standard deviations. Level of significance: * = p ≤ .05; ** = ≤ .01.

Subject x Condition interaction in the group analysis (F [7,71] = 2.87; p = .011). There does not appear to be an explanation for why S2's performance was poorer with the experimental MAP, particularly since he reported a strong preference for using this MAP in everyday life.

DISCUSSION

Subjects' reports of increased sound clarity with the experimental MAP, their preference for it particularly in difficult listening situations, and the choice of this MAP for continued use at the end of the study by seven of the eight subjects lend strong support for the benefit they received in everyday life. Analyses of subjects' responses to closed- and open-set speech tests, as well as a synthetic vowel discrimination test, provide some insight into why they perceived more benefit from the experimental than the default MAP. This benefit appears related to a combination of two factors. First, increasing the number of electrodes assigned to either the F1 or F2/F3 frequency regions by eliminating higher frequencies between approximately 6000 and 8000 Hz was associated with a small, but significant, improvement in group performance on the closed-set vowel and synthetic vowel tests with the experimental MAP. The fact that these tests were presented at relatively low levels did not preclude obtaining these significant results. Second, the elimination of high-frequency energy above 6000 Hz with the experimental MAP did not significantly reduce the information transmitted for the frication feature in the closed-set consonant test or the initial and final consonants in the open-set CNC word test. In view of the relatively low group mean scores, the low presentation level used for the closed-set consonant test may have limited the audibility of some consonants (i.e., voiceless stop, /p/; fricatives, /f/, /v/, /s/, /z/; and the nasal, /m/) with both MAPs. It is assumed that audibility was limited because, in contrast, the group mean scores for the more intense fricatives /ʃ/ and /ʒ/ alone were between 67% and 94%, a result consistent with these consonants being audible. Although the transmission of the frication feature clearly increased from 33% to ~57.5% with increasing level between 40 and 60 dB SPL, there was no significant difference between the two MAPs within a level and test material.

In contrast to the significantly higher group mean scores on the closed-set vowel and synthetic vowel tests, there were no significant differences in scores on the closed-set consonant or the CNC word tests. In addition, there were no significant differences in the information transmitted percentages for the F1, F2, or F2 movement features of medial vowels in the CNC words or the closed-set vowel test for the two MAPs. The fact that these test conditions revealed no differences between the two MAPs may be related to several factors. Based on responses to the questionnaire, subjects preferred the experimental MAP more often in difficult, rather than easy listening, situations. When cochlear implant recipients listen to speech, it is easier for them to recognize words when they can focus on the speech of a single talker rather than on multiple talkers, especially when that talker speaks clearly (i.e., clearly enunciated and a little slower than normal). It may have been that the recorded CNC word test used in this study was not difficult enough, given that it uses one male speaker who spoke clearly. Although there were two talkers of one gender presented in one test list for the closed-set consonant tests, these talkers' tokens were intentionally chosen to sound similar. Finally, none of these tests was presented in noise, a listening situation in which most subjects found the experimental MAP preferable to the default MAP.

In future research, selection of different speech stimuli may provide greater insight into the benefit of manipulating frequency boundary assignments to channels, for example, use of talkers with less well-enunciated, faster styles of speaking and multiple talkers whose vowel utterances represent a greater range of fundamental, first and second formant values (e.g., Peterson and Barney, 1952; Hillenbrand et al, 1995). Another approach would be the use of words and/or sentences presented in noise. A more quantitatively based approach to evaluating sound quality differences could be undertaken. Finally, the synthetic vowel test could be redesigned to specifically test the first and second formant discrimination affected by the changes in frequency boundaries.

An additional factor that may be related to subjective preference of the experimental MAP is the difference in the pitch ranges stimulated with the default and experimental MAPs. Examination of Figure 1 indicates that all channels coding frequencies above about 1000 Hz in the experimental MAPs are shifted to a more basal location in the cochlea by one or two electrodes relative to the default MAPs. For any given channel in this range then, the elicited pitch may be somewhat higher with the experimental MAP than with the default MAP. Recent research using CT scans suggests the pitch range elicited by cochlear implants to be quite variable with the pitch elicited by the most apical electrode estimated to be in a range between 300 and 3700 Hz (Skinner, Ketten, et al, 2002). However, this change could have countered our attempts at increasing formant resolution with an additional electrode by effectively increasing the pitch-to-channel mismatch. If this were the case, however, it seems unlikely that subjects would prefer it. Does the subjective preference for the experimental MAP suggest a more salient pitch-to-channel match? Past research (e.g., Fu and Shannon, 1999a) suggests that CI users can tolerate up to about 3 mm frequency allocation shifts. Within this 3 mm range, best performance results from allocations closest to the clinical settings with which they have grown accustomed. Although the frequency shifts between the default and experimental MAPs in the present study were relatively small, most subjects noticed a greater clarity of sound with the experimental MAP from the first day of use. Future work in this area might include radiographic analysis of electrode location in the cochlea to better understand the relation between frequency-to-electrode patterns and speech recognition in everyday life.

In the present study, the fixed number of frequency assignment tables provided by the manufacturer limited the reassignment of electrodes in the experimental MAP. As a result, two subjects had one additional electrode in the F1 region, and the remaining six had one additional electrode in the F2/F3 region of their experimental MAPs. Unfortunately, choice of channel bandwidths with the SPrint™ processor is limited to multiples of the FFT bin spacing of 125 Hz (see Figure 1). In future studies, a more expansive reassignment of electrodes in the F2/F3 region using narrower channel bandwidths should yield even better reception of information that differentiates front from back vowels, as well as improved detection of F2 and F3 transitions that are main cues for place of articulation of consonants. This reassignment can be accomplished with the SPEAR3 research processor in conjunction with the Nucleus 24 internal device (Vandali et al, 2001).

A final point of consideration is the perception by six of the eight subjects that the experimental MAP was "louder" than the default MAP. Because there was no significant difference in warble-tone thresholds between MAPs except at 6000 Hz, the perceived loudness difference must be based on something other than strictly sensitivity. We hypothesize that with the default MAP, the most basal electrode was rarely stimulated, given the natural intensity and frequency of occurrence of energy in the speech signal above 6500–7000 Hz. The experimental MAP, on the other hand, should have caused this electrode to be stimulated more often, coding signal energy from about 5000 to 6000 Hz. Coupled with the frequency allocation-to-electrode shift, this may have provided a seemingly "louder" percept.

Based on subjects' strong preferences for the experimental MAP and on small, but significant, improvements in closed-set vowel recognition and synthetic vowel F2 discrimination with the experimental MAP, it is clinically prudent to give adult cochlear implant recipients an opportunity to compare an experimental MAP that assigns more electrodes in either the F1 or F2/F3 regions with the default assignment. With four program locations on the SPrint™ body-worn speech processor, it is easy for patients to compare their original MAP to a MAP with experimental frequency boundaries. Since this study was completed, almost all Nucleus 24 CI recipients at Washington University School of Medicine in St. Louis have had an opportunity to make this comparison and have chosen to use the MAP with experimental frequency allocations in everyday life. To create a MAP with experimental frequency boundaries, it is recommended that patients have at least 16 channels in their MAP and that a frequency boundary table that is two tables below the default table be selected. The highest frequency boundary should extend to approximately 6000 Hz (5800 Hz for patients who are using 16 or 17 channels).

This study was completed before the ESPrit 3G™ ear-level speech processor was commercially available. The frequency boundary tables provided by Cochlear Americas' R126 clinical software, used to program both the SPrint™ and 3G™ processors, are different for the two processors. For patients using 19 or 20 channels in their 3G™ MAP, selecting a frequency boundary table that is one or two tables below the default table provides an additional electrode in the F1 region. For patients using 17 or 18 channels in their 3G™ MAP, selecting a frequency boundary table that is one table below the default table will provide an additional electrode in the F1 region for an 18 channel MAP and an additional electrode in the F2 region for a 17 channel MAP. For 17 or 18 channel 3G™ MAPs, selecting a frequency boundary table two tables below the default table will cause the upper frequency boundary to be less than 6000 Hz. For 3G™ MAPs with fewer than 17 channels, it is suggested that the default frequency boundary table be used.

SUMMARY

The goal of this preliminary study was to determine whether the use of frequency tables, which provided additional vowel information while sacrificing some high-frequency information, were subjectively and objectively preferable to default tables used clinically. Subjectively, seven out of eight participants preferred the nonstandard tables, citing increased clarity and loudness. To validate the subjective findings, a number of objective speech perception measures was obtained. It appears clear that limiting the high-frequency response to about 6000 Hz did not significantly alter perceptual cues in this frequency region. In addition, speech perception performance was significantly higher with the experimental MAP on a closed-set vowel test and for F2 discrimination on a synthetic vowel test. Performance on other measures was not significantly different between the two MAPs. This may have been due in part to insufficiently difficult speech tasks, and/or the shift in frequency allocation tables was not substantial enough to elicit significant changes in speech perception. Clearly, further research is needed to determine which frequency boundary allocations provide patients with the best sound quality and speech recognition with all cochlear implant systems.

Acknowledgment. We are grateful to Dianne Van Tasell for the consonant recordings and James Hillenbrand for the vowel recordings. Appreciation is expressed to the eight subjects who graciously gave their time and effort to participate in this study and to the three reviewers for their useful comments.

REFERENCES

Dowell RC, Seligman PM, Blamey PJ, Clark GM. (1987) Speech perception using a two-formant 22-electrode cochlear prosthesis in quiet and in noise. *Acta Otolaryngol* 104:439–446.

French NR, Steinberg JC. (1947) Factors governing the intelligibility of speech sounds. *J Acoust Soc Am* 19:90–119.

Fu Q-J, Shannon RV. (1999a) Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing. *J Acoust Soc Am* 105:1889–1900.

Fu Q-J, Shannon RV. (1999b) Effects of electrode location and spacing on phoneme recognition with the Nucleus-22 cochlear implant. *Ear Hear* 20:321–331.

Fu Q-J, Shannon RV. (1999c) Effects of electrode configuration and frequency allocation on vowel recognition with the Nucleus-22 cochlear implant. *Ear Hear* 20:332–344.

Harnsberger JD, Svirsky MA, Kaiser AR, Pisoni DB, Wright R, Meyer TA. (2001) Perceptual "vowel spaces" of cochlear implant users: implications for the study of auditory adaptation to spectral shift. *J Acoust Soc Am* 109:2135–2145.

Hawks JW, Miller JD. (1995) A formant bandwidth estimation procedure for vowel synthesis. *J Acoust Soc Am* 97:1343–1344.

Hillenbrand J, Getty LA, Clark MJ, Wheeler K. (1995) Acoustic characteristics of American English vowels. *J Acoust Soc Am* 97:3099–3111.

Holden LK, Skinner MW, Holden TA, Demorest ME. (2002) Effects of stimulation rate with the Nucleus 24 ACE speech coding strategy. *Ear Hear* 23:463–476.

Klatt D, Klatt LC. (1990) Analysis, synthesis, and perception of voice quality variations among female and male talkers. *J Acoust Soc Am* 87:820–857.

Luxford W, Ad Hoc Subcommittee. (2001) Minimum speech test battery for postlinguistically deafened adult cochlear implant patients. *Otolaryngol Head Neck Surg* 124:125–126.

Miller GA, Nicely PE. (1955) Analysis of perceptual confusions among some English consonants. *J Acoust Soc Am* 27:338–352.

Mueller HG, Killion MC. (1990) An easy method for calculating the articulation index. *Hear J* 49:1–4.

Nearey T. (1989) Static, dynamic, and relational properties in vowel perception. *J Acoust Soc Am* 85:2088–2113.

Parkinson AJ, Arcaroli J, Staller SJ, Arndt PL, Cosgriff A, Ebinger K. (2002) The Nucleus® 24 Contour™ cochlear implant system: adult clinical trial results. *Ear Hear* 23(suppl. 1):41S–48S.

Pascoe DP. (1975) Frequency responses of hearing aids and their effects on the speech perception of hearing-impaired subjects. *Ann Otol Rhinol Laryngol* 84(suppl. 23):1–40.

Pavlovic CV. (1986) Articulation index predictions of speech intelligibility in hearing aid selection. *ASHA* 30:63–65.

Pavlovic CV, Studebaker GA, Sherbecoe RL. (1985) An articulation index based procedure for predicting the speech recognition performance of hearing-impaired individuals. *J Acoust Soc Am* 80:50–57.

Peterson G, Barney H. (1952) Control methods used in a study of vowels. *J Acoust Soc Am* 24:175–184.

Peterson GE, Lehiste I. (1962) Revised CNC lists for auditory tests. *J Speech Hear Disord* 27:62–70.

Skinner MW. (1988) *Hearing Aid Evaluation*. Englewood Cliffs, NJ: Prentice Hall.

Skinner MW, Arndt PL, Staller SJ. (2002) Nucleus 24 Advanced Encoder conversion study: performance versus preference. *Ear Hear* 23(suppl. 1):2S–17S.

Skinner MW, Clark GM, Whitford LA, et al. (1994) Evaluation of a new spectral peak coding strategy for the Nucleus 22 channel cochlear implant system. *Am J Otol* 15(suppl. 2):15–27.

Skinner MW, Holden LK, Holden TA. (1995) Effect of frequency boundary assignment on speech recognition with the SPEAK speech-coding strategy. *Ann Otol Rhinol Laryngol* 104(suppl. 166):307–311.

Skinner MW, Holden LK, Holden TA. (1997) Parameter selection to optimize speech recognition with the Nucleus implant. *Otolaryngol Head Neck Surg* 117:188–195.

Skinner MW, Holden LK, Holden TA, Demorest ME. (1999) Comparison of two methods for selecting minimum stimulation levels used in programming the Nucleus 22 cochlear implant. *J Speech Lang Hear Res* 42:814–828.

Skinner MW, Holden LK, Holden TA, Dowell RC, Seligman PM, Brimacombe JA, Beiter AL. (1991) Performance of postlinguistically deaf adults with the wearable speech processor (WSP III) and mini speech processor (MSP) of the Nucleus multi-electrode cochlear implant. *Ear Hear* 12:3–22.

Skinner MW, Ketten DR, Holden LK, Harding GW, Smith PG, Gates GA, Neely JG, Kletzker GR, Brunsden B, Blocker B. (2002) CT-derived estimation of cochlear morphology and electrode array position in relation to word recognition in Nucleus-22 recipients. *J Assoc Res Otolaryngol* 03:332–350.

Sun JC, Skinner MW, Liu SY, Wang FNM, Huang TS, Lin T. (1998) Optimization of speech processor fitting strategies for Chinese-speaking cochlear implantees. *Laryngoscope* 108:560–568.

Välimaa TT, Määttä TK, Löppönen HJ, Sorri MJ. (2002) Phoneme recognition and confusions with multichannel cochlear implants: vowels. *J Speech Lang Hear Res* 45:1039–1034.

Vandali AE, Harrison JM, Van Hoesel R, McDermott H, Zakis J, Cowan RS, Saunders E. (2001) An advanced speech processor for cochlear implant and hearing aid research. Abstract. In: *2001 Conference on Implantable Auditory Prostheses.* 120.

Vandali AE, Whitford LA, Plant KL, Clark GM. (2000) Speech perception as a function of electrical stimulation rate: using the Nucleus 24 cochlear implant system. *Ear Hear* 21:608–624.

Van Tasell DJ, Greenfield DB, Logemann JJ, Nelson DA. (1992) Temporal cues for consonant recognition: training, talker generalization, and use in evaluation of cochlear implants. *J Acoust Soc Am* 92:1247–1257.

Walker G, Dillon H, Byrne D. (1984) Sound-field audiometry: recommended stimuli and procedures. *Ear Hear* 5:13–21.

Wang MD, Bilger RC. (1973) Consonant confusions in noise: a study of perceptual features. *J Acoust Soc Am* 54:1248–1256.

Wilson BS, Lawson DT, Finley CC, Wolford RD. (1995) New processing strategies in cochlear implantation. *Am J Otol* 16:669–675.