Interactions of Hearing Aid Compression Release Time and Fitting Formula: Effects on Speech Acoustics

John C. Ellison
Frances P. Harris
Thomas Muller

Abstract
The effects of the interaction of compression release time and prescribed gain on running speech processed through a hearing aid on KEMAR was investigated. A digital instrument was programmed to fit a mild to moderate sloping hearing loss using probe microphone measures to reach targets prescribed by NAL-NL1, DSL I/O, FIG.6 or ASA2p with release times of 40 and 640 ms for each condition. Recordings were made through KEMAR and analyzed to determine the long-term-average-speech spectra, consonant-to-vowel ratios and the RMS amplitude of 32 phonemic units. Aided and unaided results were compared. Within each prescriptive formula, changes in release time affected all of the speech measures subsequent to programming the instrument to a static-composite signal. The short release-time condition produced the greatest alteration to the speech signal. Release time may need consideration when fitting hearing aids to target gain prescriptions.

Key Words: Compression, compression attack time, compression release time, compression ratio, compression threshold, wide dynamic range compression, prescriptive formula, Knowles Electronics Manikin for Acoustic Research, speech envelope.

Abbreviations: ASA2p = Adaptive Speech Alignment, B&K = Brüel and Kjær, CVR = Consonant-Vowel Amplitude Ratio, DSL I/O = Desired Sensation Level Input / Output, KEMAR = Knowles Electronics Manikin for Acoustic Research, LTASS = Long Term Average Speech Spectrum, NAL-NL1 = National Acoustic Laboratories nonlinear procedure Version 1, REIG = Real Ear Insertion Gain, REUG = Real Ear Unaided Gain, RMS = Root Mean Square, SPL = Sound Pressure Level, VCV = Vowel-Consonant-Vowel Stimuli, WDRC = Wide Dynamic Range Compression

Sumario:
Se investigaron los efectos de la integración del tiempo de liberación de la compresión (compression release time) y de la ganancia prescrita, utilizando lenguaje corrido a través de un auxiliar auditivo en el KEMAR. Un dispositivo digital fue programado para amplificar una hipoacusia leve a moderada con pendiente progresiva, utilizando mediciones de un micrófono de prueba que alcanzaran los objetivos indicados en NAL-NL1, DSL I/O, FIG.6, ASA2P, con tiempos de liberación de 40 a 640 ms para cada condición. Los registros se realizaron a través del KEMAR y fueron analizados para determinar los espectros promedio a largo plazo del lenguaje, las tasas consonante/vocal y la amplitud RMS de 32 unidades fonémicas. Se compararon los resultados con y sin amplificación. Dentro de cada fórmula de prescripción, los cambios en el tiempo de liberación afectaron todas las medidas de lenguaje, luego de programar el instrumento para una señal estática-compuesta. La condición de tiempo de liberación corto produjo la mayor alteración de la señal de lenguaje. El tiempo de liberación debería ser tomado en cuenta cuando se ajuste un auxiliar auditivo para alcanzar una ganancia de prescripción.

*University of Arizona, Department of Speech and Hearing Sciences, Tucson, AZ
Reprint Requests: Frances P. Harris, Speech & Hearing Sciences, 1131 E. 2nd St., Tucson, AZ 85721
Phone: (520) 626-7530 Email: fpharris@email.arizona.edu

Portions of this paper were presented at the American Academy of Audiology 2002 convention; April 20, 2002; Pennsylvania Convention Center, One Convention Center Pl., 1101 Arch St., Philadelphia, PA 19107
Although selection of an appropriate fitting formula and of compression characteristics are both routinely considered when fitting a hearing aid to an individual, the interactions of the two are often not. Gain as a function of frequency may be established by one of a number of prescriptive formulas, none of which include compression release time in their calculations. Yet, changes in compression release time are known to influence speech. It would seem appropriate to know how gain may be influenced by changes in compression release time and what effects such changes might have on speech. Such interactions are the topic of this investigation.

Compression can alter speech characteristics in several ways. Compression-release time can affect the temporal amplitude-modulation envelope of a speech signal (Van Tasell, 1993). The effect of long release times on the modulation envelope is minimal relative to short release times. Kuk (1998) claims that long release times maintain the short-term intensity relationships in speech. Short release times, those equivalent to or shorter than the duration of a syllable, allow the compression to vary during the course of a word. As such, wide dynamic range compression (WDRC) instruments with short release times compress high-level low-frequency segments (i.e., vowels) and release fast enough to provide gain to low-level high-frequency segments (i.e., consonants) (Van Tasell, 1993). Relative to the input, the amplitude envelope of speech will then be smoothed or distorted and the consonant-vowel amplitude ratio (CVR) increased (i.e., consonant amplitude is increased) at the output.

Rosen (1992) described the speech envelope as amplitude fluctuations in speech between 2 and 50 Hz that provide segmental cues for manner of articulation, voicing, vowel identity, and prosody. Because individuals with profound hearing impairments typically have degraded auditory frequency selectivity, Rosen implies that these individuals rely greatly upon temporal cues, such as the amplitude envelope of speech. In support of this, several investigators have indicated a positive correlation between the amount of access to speech envelope cues and performance on word or syllable recognition tasks for subjects with profound hearing loss (Erber, 1972; Boothroyd et al, 1988).

Several investigators (Gordon-Salant, 1986, 1987; Freyman and Nerbonne, 1989; Preves et al, 1991; Balakrishnan et al, 1996; Souza and Turner, 1996; Hickson and Byrne, 1997; Sammeth et al, 1999; Smith and Levitt, 1999; Souza, 2000;) have studied how the amplitude envelope or CVR of speech affects the intelligibility of speech units (i.e., items no longer than a word); however, results have been conflicting. Outcomes obtained using longer stimuli are more consistent. Souza and Kitch (2001) studied the effects of altering amplitude envelope cues on sentence identification while controlling for audibility for normal-hearing and hearing-impaired adults each divided into groups of elderly and young adults. The authors determined that recognition scores became poorer as compression ratios increased from 1:1 to 2:1 and from 2:1 to 5:1, thus demonstrating a negative correlation between performance and the degree to which amplitude envelope cues were altered.

Van Tasell and Trine (1996) conducted three experiments to study the degree to which the amplitude envelope contributes to VCV disyllabic and sentence recognition in normal-hearing subjects. The authors found that severe compression did not result in significantly poorer performance on VCV identification relative to uncompressed conditions. Subjects were still able to classify disyllabic stimuli on the basis of compressed envelope cues after spectral and periodicity cues were removed. Importantly, the authors found that compression of envelope information and not the removal of periodicity information adversely affected subject perform-

**Palabras Clave:** Compresión, tiempo de ataque de compresión, tiempo de liberación de la compresión, tasa de compresión, umbral de compresión, compresión del rango dinámico amplio, fórmula de prescripción, KEMAR: Maniquí Electrónico Knowles Para Investigación Acústica, envolvente del lenguaje.

**Abreviaturas:** ASA2P = Alineamiento adaptativo del habla; B&K = Brüel y Kjaer; CVR = Tasa de amplitud consonante/vocal; DSL I/O = Entrada / salida del nivel deseado de sensación; KEMAR = Maniquí Electrónico Knowles Para Investigación Acústica; LTASS = Espectro promedio del lenguaje a largo plazo; NALNL1 = Versión 1 del procedimiento no-lineal de los Laboratorios Nacionales de Acústica; REUG = Ganancia no amplificada en Oído Real; RMS = Raíz cuadrada media; SPL = Nivel de presión sonora; VCV = Estímulo vocal-consonante-vocal; WDRC = Compresión de rango dinámico amplio.
ance for sentence recognition after spectral information had been removed. Van Tasell and Trine concluded that, at least for normal-hearing subjects, envelope cues are needed for sentence identification. Further evidence supporting envelope cues in sentence recognition comes from Drullman (1995), who studied the role of envelope and fine structure cues for sentence recognition in normal-hearing subjects. He found that the sentences with intact envelope cues and random fine structure cues were intelligible; whereas, the sentences with intact fine structure cues but random envelope cues yielded mean recognition scores of 17%. Overall, studies that incorporate sentences demonstrate the importance of envelope cues for sentence recognition.

During the fitting process, hearing aids are set to gain requirements specified by prescriptive formula. Three common prescriptive formulas used to fit nonlinear instruments include Desired Sensation Level Input/Output (DSL I/O; Cornelisse et al, 1994), FIG.6 (Killion and Fikret-Pasa, 1993), and National Acoustic Laboratories’ nonlinear procedure, Version 1 (NAL-NL1; Byrne et al, 2001). These formulas share the common goal of determining gain requirements based on the level of the input signal. Additionally, hearing aid manufacturers often incorporate their own prescriptive formulas for fitting nonlinear hearing aids.

The practice of fitting all hearing aids to a single prescriptive target may not be advisable. For example, Kuk and Ludvigsen (1999) claim that no prescriptive formula accounts for attack or release times. This becomes problematic because two hearing aids with different response times may respond differently in real-world situations under compression even if both hearing aids are matched to the same prescriptive formula. Kuk and Ludvigsen claim that the real-world dynamic outputs of hearing aids with similar short attack and release times approximate the predicted static input/output curve (i.e., the response assumed by prescriptive formulas). Conversely, they claim that if the attack time were substantially shorter than the release time, then the average real-world dynamic output of the hearing aid would be below that assumed by the static or prescription curve. They assert that matching prescriptive targets is acceptable with fast-acting WDRC instruments, but not for slower-acting hearing aids. Thus, Kuk and Ludvigsen question the value in matching prescriptive targets in hearing aids with longer release times.

Given that compression circuits are present in most hearing aids and that prescriptive methods are used routinely during hearing aid fittings, the nature of their interactions should be clarified. This study investigated the influence that variations in specific compression parameters produced on a dynamic speech signal following adjustment of the gain of a hearing aid according to one of four prescriptive formulas designed for nonlinear hearing aids. Based on previous literature, it was expected that a short release time would smooth the amplitude envelope of speech relative to a long release time. Also, based on claims by Kuk and Ludvigsen (1999), it was expected that the hearing aid would respond to a dynamic speech signal with relatively greater output under a short release-time than a long release-time condition, even if the hearing aid response matched the same target for both release times. Differences across prescriptive formulas would be expected because of inherent differences in prescribed gain and compression ratio. This study addressed two questions:

1. For a prescribed setting, do changes in compression release time differentially affect the long-term-average-speech spectrum (LTASS), the amplitude envelope, and phoneme amplitude of a three-sentence speech signal?
2. Which compression release time has the greatest effect on a speech signal for each prescriptive condition?

**METHOD**

Pre-recorded running speech was presented in the soundfield and received by a Knowles Electronics Manikin for Acoustic Research (KEMAR) in one unaided and four aided conditions. Responses were recorded and stored for analysis offline. Repeated measures were made of the recording conditions to assess measurement error. To determine the overall and long-term output of the hearing aid, the LTASS was measured in each condition. Phonemic output was measured to observe general phonemic trends and to corroborate LTASS results. For this, absolute level (in dB SPL) was measured for each phoneme in the speech signal. Finally, CVRs were measured to assess the degree of alteration in the amplitude envelope.
Speech Signal

Running speech was selected to allow measurement of the effect of compression on a dynamic signal. The signal was pre-recorded and consisted of three sentences:

1. Peter will keep at the tall chilly peak (pitə ^ wɪlk æ/t\l/ tʃIli pik).
2. The blue spot is on the thin key again (θə blu spɒt ɪzn /θ\n ki ægcn).
3. When he comes home we’ll feed him fish (wɛn /hɪ/ kæms həmil /hI/m fj).

These sentences were chosen for three reasons. First, they include most phonemes found in American English, including all voiceless consonants and all vowels except diphthongs and /e, 3, U, 3^, a/. An adequate variety of vowels was needed to measure relative changes in their amplitude and to calculate the CVR. Voiceless consonants were needed to exclude voicing in the contrast between vowels and consonants to facilitate calculation of CVRs. Second, vowels and voiceless consonants are represented in the initial, medial, and final position. Third, these sentences represent running speech that a listener might encounter in a real-world situation.

The sentences were spoken by a male talker and were recorded using a microphone (Sennheiser 825S) and a microphone amplifier (Tucker-Davis Technologies MA2) with the gain set to 35 dB. The signal was routed to a sound card (Sound Fusion Wave) and was recorded using speech recording and analysis software (Praat 3.8.69) at a monaural 44100 Hz sample rate.

During the recording, the talker repeated each sentence three times. From this sample, the middle sentence was selected and placed onto one final recording. Sentences were separated by 400 ms of silence. The recording was converted to a 16-bit WAV file. On the recording, a 1000 Hz marker tone of 140.7-ms duration and 19 dB below the overall level of the recording was placed before the onset of the first sentence to mark the beginning of the recording.

Hearing Aid Description

An Oticon Digifocus II Power behind-the-ear digital hearing aid was used because its release time can be modified. This hearing aid contains two channels with three frequency bands in the low-frequency channel and four frequency bands in the high-frequency channel (cutoff frequency = 1500 Hz). Other than release time, all parameters remained constant across measurement conditions. The compression release time was either 40 or 640 ms for each measurement condition and was kept equal for both hearing aid channels. These release times represent the extreme settings within the hearing aid and are within the range of release times implemented in commonly used hearing aids currently on the market.

Prescriptive Formulas and Hearing Loss Configuration

Four nonlinear prescriptive fitting formulas were used to set the gain of the hearing aid: DSL I/O (Cornelisse et al, 1994), FIG.6. (Killion and Fikret-Pasa, 1993), NAL-NL1 (Byrne et al, 2001), and the Adaptive Speech Alignment (ASA^p) algorithm. The former three represent common nonlinear prescriptive formulas that recommend fitting to the static input/output gain curve (Kuk, 2001). Digital Oticon hearing aids incorporate the ASA^p. Oticon describes ASA^p as an algorithm that compensates for recruitment and controls for the upward spread of masking by incorporating syllabic compression in the low frequency channel, and attempts to achieve comfort and speech intelligibility by adapting the gain for changes in

| Table 1. Prescribed Gain (in dB SPL) From 250 to 4000 Hz Using a 65 dB SPL Input For Each Prescriptive Formula and Selected Compresssion Ratio In Each Hearing Aid Channel. |
|-----------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
|                  | 250       | 500       | 750       | 1000      | 2000      | 3000      | 4000      | Compression Ratio |
|                  | Low Freq. | High Freq.|
| ASA^p            | 1.6:1     | 2.3:1     |
| DSL I/O          | 1.4:1     | 2.1:1     |
| FIG.6            | 1.8:1     | 5:1       |
| NAL-NL1          | 1.7:1     | 2.3:1     |
input in the high frequency channel. The rationale is to place speech into a comfortable listening range (Bentler and Duve, 2000) and maintain the spectro-temporal detail of speech in the high frequencies. Because the Oticon Digifocus II hearing aid is designed for the ASA2p algorithm, ASA2p was included. Although a variety of instruments might be fit using these formulas, none take into account the range of possible release times available. If prescriptive formulas attempt to approximate the predicted static input/output curve, then prescriptive formulas, including NAL-NL1 and DSL I/O, must assume fast release times. In contrast, because FIG.6 was developed for use with the K-amp circuit, it must assume variable release times. ASA2p incorporates time constants described by the algorithm.

NOAH 2.0a was employed to display prescriptive targets based on a hearing loss that resembles a typical audiometric pattern of the older adult population aged in their mid 70’s (Harford and Dodds, 1982). This configuration (see Figure 1) was used for all conditions. Table 1 displays the prescribed gain between 250 and 4000 Hz for a 65 dB SPL input for each fitting formula as well as the selected compression ratio in each channel that best approximated target. Table 1 does not include frequencies above 4000 Hz because of the limited response of the hearing aid and inability to manipulate gain above this frequency.

When the hearing aid was set to the ASA2p algorithm, attack and release times were changed to equate the settings used for the three prescriptive formula conditions. Specifically, the attack time was adjusted from 20 to 5 ms, and the release time was adjusted from 80 ms in the low-frequency channel and 320 ms in the high-frequency channel to 40 or 640 ms in both channels. These changes did not alter the prescribed gain for the hearing instrument but may have altered the response of the hearing instrument from what the manufacturer intended.

**Hearing Aid Programming**

The hearing aid was programmed to approximate multiple input targets visually, which is a common clinical verification procedure. Matching to multiple input targets was intended to account for the changing gain requirements of nonlinear formulas at various input levels. OtiSet 4.20 fitting software was used to program and manipulate the electro-acoustic parameters of the hearing aid. NOAH 2.0a served as the software platform for OtiSet 4.20 and was utilized to observe the real-time response of the hearing aid while it was programmed to target. Programming the hearing aid to the ASA2p algorithm was accomplished in OtiSet by choosing the “calculate prescribed settings” option on the OtiSet toolbar without real-ear verification.

**Probe-microphone measures.** Real-ear probe microphone measures were obtained to visually approximate the real-ear insertion gain (REIG) of the hearing aid to prescriptive targets. The real-ear unaided gain (REUG) was obtained so that the REIG could be determined. A modulated white-noise signal generated by a Madsen Aurical HiPro console and presented through the Aurical HiPro speaker served as the probe microphone signal used to match the response of the hearing aid to inputs of 40, 65, 70, and 90 dB SPL. The hearing aid gain in seven frequency bands (i.e., 250, 750, 1000, 2000, 3000, 4000, and 5000 Hz) was adjusted to adhere to the best possible visual approximation of target gain. The process of visually matching hearing aid gain to target gain verified that no visually detectable deviation occurred subsequent to release time manipulations within each prescriptive condition. All programming modifications performed in OtiSet were routed to the hearing aid through the Aurical HiPro console and Oticon programming cable.

**Instrumentation and Calibration**

All data recordings were made from KEMAR positioned in a sound booth at a 0-degree azimuth, 1
From a loudspeaker, and level with the head of KEMAR. The loudspeaker has a 6.5-in. woofer and 1-in. tweeter, a response of 57 Hz to 25 kHz, a crossover of 2.2 kHz (6 dB/octave high-pass and 12 dB/octave low-pass), and a sensitivity of 2.83V at 1 m. The loudspeaker was powered by a 6-channel power amplifier (Rotel, model RB-976).

Using Praat 3.9.22 software (Boersma and Weenink, 2001), a 10-s 1000 Hz calibration tone was created to set the VU meter of the audiometer. The level of the calibration tone was set to the peak of the root-mean-square (RMS) envelope of the speech signal (excluding the 400-ms silent intervals).

A Type I sound level meter (Brüel and Kjær [B&K], model 2204) and 1-in. condenser microphone (B&K, model 4145) were used to calibrate the signal in the sound field. The level for the calibration tone measured at the ear of KEMAR in the soundfield was 76 dB SPL. The range of the level of the speech signal corresponding to this setting was 65 to 77 dB SPL, and is within the range of normal conversational level speech when the speaker's own voice is accounted for (Skinner, 1988). Because the hearing aid compression kneepoint was fixed at 50 dB SPL, this level ensured that the hearing aid would be in compression during signal presentation.

The sound level meter and microphone were checked for calibration at 110 dB SPL and 94 dB SPL using a Quest Electronics calibrator (model CA-12B) and a B&K acoustical calibrator (model 4231), respectively. The real-ear probe microphone was calibrated with a 65 dB SPL swept pure tone.

**Signal Delivery and Recording**

The signal was presented via media software (Windows Media Player, version 6.01.05.0217) on a computer (Hewlett Packard Pentium III) and routed to the loudspeaker through an Aurical HiPro console. The signal level was controlled using an audiometer (Madsen Aurical).

The signal was recorded via a 1/2-in. microphone (B&K, type 4192), which was coupled to the Zwislocki coupler of KEMAR. A flexible adapter (model UA0122) coupled the 1/2-in. microphone with a preamplifier (B&K, model 2669C). The signal was then led to a signal conditioner (B&K Nexus, model ZX2690) and routed into a programmable attenuator (Tucker-Davis PWS 25). The attenuator was set to 0 dB under unaided conditions and to 24 dB under aided conditions. These settings permitted an optimized signal for computer processing. From the attenuator, the signal was routed to a sound card (Sound Fusion Wave) situated in a computer (Hewlett Packard Pentium III). All signals were recorded at a sampling rate of 44100 Hz and stored in WAV format on one channel.

**Reference tone.** Because the level calculation within the Praat software is based on an arbitrary reference, a reference tone with a known level (94 dB SPL, B&K calibrator 4231) was recorded to convert phoneme levels from speech signal data recordings into absolute levels. To determine the absolute level of phonemes within a data recording, the difference between the 94 dB SPL calibration signal and the recorded level observed in Praat was added to the RMS calculation for each phoneme.

**Estimation of measurement error across data recordings.** To estimate the measurement error expected from a routine fitting, each condition was recorded on five separate occasions.

**Procedure**

The audiogram was entered into NOAH 2.0a. The sound level meter and microphone were checked for calibration. The condenser microphone of the sound level meter was hung from the ceiling so that it rested at the ear-level of KEMAR. The soundfield and the VU meter of the audiometer were calibrated. The 94 dB SPL reference tone was recorded. The B&K 1/2-in. microphone was then situated in KEMAR. The probe microphone was calibrated and placed in the ear canal of KEMAR. Using a modulated white-noise signal, a real-ear unaided response curve was obtained. The attenuator was set according to the signal condition. In unaided (i.e., baseline) conditions, two sequences of the speech signal were presented, recorded, and saved. In aided conditions, the hearing aid was placed on the right ear of KEMAR and connected to the HiPro console. The hearing aid was programmed while set to a release time of 640 ms. The speech signal was presented, recorded, and saved as in the unaided conditions. Compression release time was adjusted to 40 ms and the signal was again presented, recorded, and the recording stored for later analysis. This procedure was repeated on five separate occasions so that each condition was recorded five times.
Data Analysis

Data preparation. The original speech recording was phonemically segmented into 68 units. To compare the level of phonemes across all data recordings, phonemic time segments determined from the original speech signal recording were matched across all data recordings. This was accomplished by representing the onset of the marker tone (from the original speech signal and subsequent data recordings) as the beginning point of all data recordings. The area following the last phoneme was removed on all recordings to establish a common endpoint.

MATLAB 6.1.0.450 was used to calculate RMS values within each phonemic time segment and the LTASS in one-third octave bands between 250 and 5000 Hz for each data set. This frequency range represented the frequency bands that are adjustable within the hearing aid.

Measures and analysis. Measures for analysis included: LTASS in relative dB, CVR in dB, and the phoneme amplitudes in dB SPL.

From the determined RMS levels, the phoneme amplitudes in each segment of the data recordings were converted to dB SPL. The level for stops was based on burst and aspiration duration (Freyman and Nerbonne, 1989; Hickson and Byrne, 1997). Each condition was compared to baseline and other aided conditions as a function of release-time variation. The first phoneme from all recordings was removed in the analysis due to minimal compression effects. All data sets were equivalently balanced, resulting in 32 total phonemic units.

The CVR was determined in dB by subtracting the consonant from the vowel dB SPL to minimize negative values. Only voiceless consonant-vowel contrasts were used for these calculations because they represent the greatest contrasts among speech elements. An average CVR was determined in each condition and used for comparisons between conditions and to baseline as a function of release-time variations. The first consonant-vowel contrast in all data recordings was omitted due to minimal compression effects. All data sets were equivalently balanced, resulting in 18 consonant-vowel contrasts.

The LTASS is expressed in relative dB and represents the energy in 14 one-third octave bands. A common reference of 0.000000000001 was used to convert RMS to dB values. This reference was used to avoid results less than 1.0 and hence, negative dB values. Relative dB results were derived and used for comparisons between conditions and to baseline.

Statistical analysis. A multiple comparison ANOVA was performed for all measures. The influence of four variables (i.e., frequency, prescriptive formula, release time, and trial) and their interactions upon the LTASS were assessed. Also, four variables (i.e., phoneme, prescriptive formula, release time, and trial) and their interactions upon the CVR and phoneme level were assessed. An alpha of 0.02 was used to adjust for multiple comparisons. A Tukey Studentized-Range (HSD) post-hoc analysis (α=0.02) was performed for all measures.

Generalizability theory was applied to all measures to predict the proportion of expected variance of each of the independent component variables discussed for ANOVA upon their respective dependent variables.

For descriptive, ANOVA, and Generalizability theory analysis, CVR and phoneme amplitude data sets were balanced by randomly removing repeated observations and applying changes uniformly across trial recordings. This created data sets with homogeneity of variance and allowed direct comparison between all statistical analyses to be made.

RESULTS

Long Term Average Speech Spectrum (LTASS)

The effect of changes in release time on the LTASS was determined by calculating the difference in dB between each one-third-octave band measured in the 40- and 640-ms release-time conditions. Figure 2 displays these mean differences with the hearing aid set to the ASA2p, DSL I/O, FIG.6, and NAL-NL1 algorithms, respectively. Relative to the 640-ms conditions, the mean levels for all of the 40-ms conditions were greater: ASA2p = 4.32 dB (sd = 0.17), DSL I/O = 3.21 dB (sd = 0.56), FIG.6 = 5.47 dB (sd = 1.05), and NAL-NL1 = 2.8 dB (sd = 0.61). FIG.6 was the most affected by changes in release time; whereas, NAL-NL1 was the least affected. In all prescriptive conditions, the one-third octave bands with center frequencies of 1600, 1250, and 2000 Hz were most altered with changes in release time. The 250, 315, and 630 Hz bands were least affected by changes in release time under all formula conditions, except for DSL I/O in which the 5000 Hz band was the least sensitive; however, this applied to only one frequency. Otherwise, findings were compatible with the other three formulae.

The effect of each release time on the unaided LTASS in each formula was determined by calculating the difference (dB) in the LTASS between aided and unaided conditions (i.e., gain). Because the hearing aid used a cutoff frequency of 1500 Hz, the mean gain from one-
third octave bands was assessed within a high- and low-frequency channel using a cutoff frequency of 1500 Hz. Results suggest that within each prescriptive condition, the hearing aid provided more gain when it was set to a 40- rather than a 640-ms release time. With the exception of measures within DSL I/O, the aided LTASS in the low-frequency channel was less than the unaided LTASS under all conditions (i.e., negative gain). This was likely the result of a combination of factors. First, minimal gain was prescribed in the low-frequency channel; second, the signal would have activated the hearing aid's compression in the low-frequency channel and thereby reduced the gain in that channel; third, the occluding ear mold attenuated the signal reaching the ear canal of KEMAR. The greatest mean gain across all 14 one-third octave bands occurred when the hearing aid was set to DSL I/O with a 40-ms release time. Due to the degree of negative gain below 1500 Hz, the hearing aid provided negative overall gain (mean across all octave bands) when it was set to ASA2p, even though the gain in the high-frequency channel was still positive.

The results of the ANOVA for LTASS measures indicated that formula, release time, and frequency significantly contributed to the variance of LTASS measures. The interactions of formula and frequency, formula and release time, and release time and frequency significantly contributed to the variance. Tukey HSD post-hoc tests ($\alpha=0.02$) revealed significant effects for all formula and release-time combinations when compared to unaided conditions. The results of the Generalizability theory on LTASS measures indicated that frequency, prescriptive formula, and the interaction
between frequency and formula contributed to roughly 97% of the predicted variance estimates of LTASS measures. Release time contributed to less than 2% of the variance.

Consonant-Vowel Ratio

The effect of changes in release time on the amplitude envelope was determined by obtaining the difference (dB) in the CVR between 40- and 640-ms release-time conditions. Figure 4 displays the mean differences with the hearing aid set to the ASA²p, DSL I/O, FIG.6, and NAL-NL1 algorithms. The mean CVRs for the 40-ms conditions are less than those of the 640-ms conditions by 0.92 (sd=0.44), 1.51 (sd=0.41), 1.82 (sd=0.62), and 1.12 (sd=0.35) dB for ASA²p, DSL I/O, FIG.6, and NAL-NL1 conditions, respectively. Results suggest that the CVR is most affected by release-time changes with the hearing aid set to FIG.6. Within all formulas, /θ/ was the least affected by release-time changes. Conversely, /t/ was most affected by release-time changes within all formula conditions. Generally, contrasts that included voiceless stops were most sensitive to release-time changes.

The effect of each release time on the natural amplitude envelope within each prescriptive

Table 2. The Five Phonemes Least and Most Affected by Changes in Release Time.

<table>
<thead>
<tr>
<th>Fitting Formula</th>
<th>Least Affected</th>
<th>Most Affected</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASA²p</td>
<td>/b/ *</td>
<td>/w/ *</td>
</tr>
<tr>
<td>DSL I/O</td>
<td>/b/ *</td>
<td>/w/ *</td>
</tr>
<tr>
<td>FIG.6</td>
<td>/l/ **</td>
<td>/l/ *</td>
</tr>
<tr>
<td>NAL-NL1</td>
<td>/l/ **</td>
<td>/l/ *</td>
</tr>
</tbody>
</table>

* Phoneme represented in 3 of 4 fitting formula  
** Phoneme represented in all fitting formula
formula was determined by calculating the difference (dB) in the CVR between aided and unaided conditions. Figure 4 displays these mean differences with the hearing aid set to ASA/p, DSL I/O, FIG.6, and NAL-NL1. Relative to unaided conditions, negative dB values represent reduced differences in level between the voiceless consonant and neighboring vowel under aided conditions; whereas, positive values represent increased differences. These results suggest less alteration of the natural CVR with the hearing aid set to a 640-ms release time rather than a 40-ms release time. The greatest alteration of the natural CVR occurred when the hearing aid was set to FIG.6 and a 40-ms release time. Conversely, the least alteration of the natural CVR occurred with the hearing aid set to DSL I/O and a 640-ms release time.

The results of the ANOVA indicated that formula, release time, consonant-vowel contrast, and trial significantly contributed to the variance of the CVR. Also, the interactions of formula and release time, formula and consonant-vowel contrasts, release time and consonant-vowel contrasts, and the interactions between all three (i.e., formula, release time, and consonant-vowel contrasts) contributed to the variance of the dependent variable. Tukey HSD post-hoc tests (α=0.02) revealed significant effects for all formula and release-time combinations except for the comparison between NAL-NL1 and FIG.6 with 640-ms release times, and between NAL-NL1 and DSL I/O with 40-ms release times. All release-time effects were significant. Tukey HSD post-hoc tests (α=0.02) also revealed significant effects for all formula and release-time combinations when compared to baseline.

The results of the Generalizability theory on CVR measures indicated that the voiceless consonant-vowel contrasts contributed to the majority of predicted variance estimates of CVR measures. Release time alone contributed to approximately 2% of the estimated variance. The combined interactions between the contrasts and release times, and the contrasts and formulas contributed to approximately 10% of the predicted variance estimates.

### Phoneme Amplitude

The effect of changes in release time on phoneme amplitude was determined by calculating the difference (in dB) in phoneme amplitudes between the 40- and 640-ms conditions. Relative to the 640-ms conditions, the mean output for the 40-ms conditions was greater by 3.12 (sd = 0.4), 2.57 (sd=0.55), 4.32 (sd=0.77), 2.51 (sd=0.43) dB for ASA/p, DSL I/O, FIG.6, and NAL-NL1, respectively. Thus for the phoneme amplitudes, FIG.6 was the most affected by release-time changes whereas NAL-NL1 was least affected. This pattern confirms the LTASS results. Table 2 displays the five phonemes least and most affected by changes in release time across all prescriptive conditions. With some exceptions, stops were generally more sensitive to release-time changes than were fricatives. Under all conditions, /o/ was the vowel most affected by release-time changes. The vowel least affected by release time was /I/. Under ASA/p, FIG.6, and NAL-NL1, /b/ and /w/ constituted the only units with greater output in the 640-ms conditions than in the 40-ms conditions.
The effect of release time on the natural phoneme amplitude within each formula was determined by calculating the difference in dB (i.e., gain) between aided and unaided conditions. Figure 5 displays the mean differences with the hearing aid set to the ASA²p, DSL I/O, FIG.6, and NAL-NL1 algorithms. These results suggest that within each respective formula, the hearing aid provided more gain when the hearing aid was set to a 40- rather than a 640-ms release time. The most gain occurred with the hearing aid set to DSL I/O with a 40-ms release time. Conversely, the least gain occurred with the hearing aid set to ASA²p with a 640-ms release time. This pattern generally agrees with LTASS results.

The results of the ANOVA indicated that the formula, release time, phoneme, and trial significantly contributed to the variance of the phoneme amplitude. Also, the interactions of formula and release time, formula and phoneme, release time and phoneme, and the interactions between all three (i.e., formula, release time, and consonant-vowel contrasts) contributed to the variance of the dependent variable. Tukey HSD post-hoc tests (α=0.02) revealed significant effects for all formula and release-time combinations except for the comparison between NAL-NL1 with a 40-ms release time and FIG.6 with a 640-ms release time. All release-time effects were significant. Tukey HSD post-hoc tests (α=0.02) also revealed significant effects for all formula and release-time combinations when compared to baseline.

The results of the Generalizability theory on phoneme amplitude indicated that the phonemes, then formulas contributed the most to variance estimates, and together accounted for roughly 86% of the variance. Release time contributed to approximately 7.5% of the estimated variance. The interaction between phoneme and release time accounted for less than 4% of the estimated variance.

SUMMARY AND DISCUSSION

The purpose of this investigation was to study the interactive effects of hearing aid release time and prescribed gain on the output of a hearing aid in response to a speech signal. Results supported that within ASA²p, DSL I/O, FIG.6, and NAL-NL1, changes in release time had small but significant effects on LTASS, CVR, and phoneme amplitude. Overall, gain and output were greater with the hearing aid set to a 40-ms than a 640-ms release time. High frequencies were more affected by changes in release time than were other frequencies. Under 40-ms release time conditions, differences in level between the vowels and neighboring voiceless consonants were reduced. Phoneme amplitudes were greater with the hearing aid set to a short release time, which supported LTASS results. Stops were generally more affected by changes in release time than were fricatives. Across all measures, the effect of changes in release time was greatest for the FIG.6 prescription. These results support those of other investigators and provide the basis for some limited clinical implications.

Both Van Tasell (1993) and Kuk (1998) reported that long release times preserve the amplitude fluctuations in the speech envelope; whereas, short release times smooth the envelope. Our results confirmed this effect. Frequencies above 1500 Hz were more affected by changes in release time relative to low frequencies. Several authors (Drullman, 1995; Van Tasell and Trine, 1996; Souza and Kitch, 2001) have demonstrated the importance of envelope cues for sentence recognition. A positive correlation exists between the amount of access to speech envelope cues and performance on word or syllable tasks for subjects with profound hearing loss (Erber, 1972; Boothroyd et al., 1988,). Kuk and Ludvigsen (1999) asserted that if the attack time were substantially shorter than the release time, then gain prescribed by target would be insufficient. Because most nonlinear prescriptive formulas (with the exception of FIG.6) assume short release times, fitting long-release-time hearing aids to prescription may provide less gain and output than what is needed by the listener in the real world.

Among the prescriptive formulas used in the current study, FIG.6 prescribed the greatest compression ratio and had the most effect on the dependent variables with release time changes. This result is not unexpected: the greater the compression ratio, the larger the recovery in gain after compression. The effect would also likely be greater with more prescribed gain. For example, a nonlinear hearing aid programmed for a profound hearing loss would utilize higher compression ratios and more gain, and perform within closer proximity to its upper limit. This would likely result in greater gain, output, and speech envelope discrepancies with changes in release time. Therefore, failure to account for release time could be detrimental for speech intelligibility in individuals with profound hearing loss because of their greater reliance on envelope cues (Erber, 1972; Boothroyd et al, 1988) and the need for high levels of gain. The two-channel hearing aid used in the
study would have likely altered the speech envelope more than a single-channel device (Plomp, 1988). Such release-time effects, therefore, may be less consequential in a single-channel device.

It is unknown how the acoustic changes measured in this study would affect a listener’s performance. If variations up to 8 dB (e.g., FIG.6 at 2000 Hz) in the higher frequencies are clinically significant for speech intelligibility, then changing the release time or fitting a slow-release-time hearing aid to prescription could have adverse effects on speech intelligibility. Despite the statistically significant effect that changes in release time had on the dependent variables, such changes resulted in small absolute differences between the 40- and 640-ms conditions. Such small differences might have minimal clinical consequences. Release time accounted for a small proportion of the estimated variance relative to other independent variables (e.g., frequency, vowel-consonant contrasts, and phoneme). Therefore, despite the significant effects from changes in release time on speech, clinicians may need to be relatively more concerned with other variables. However, it is possible that if release time were ignored, then the practice of fitting hearing aids to prescriptive targets might result in poorer performance for a hearing-impaired listener. Because the practice of matching hearing aids to target is essential and generally more relied upon in the beginning of the fitting process, hearing aid acceptance could be affected, especially for first time hearing aid users. Direct evaluation of these effects would be needed to make such a determination.

There were limitations to the study that should be considered when interpreting the findings. Using averaged data, such as the LTASS, may undervalue the effect of release time on the lowest amplitude elements of speech. Also, our instrument maintained a fixed release time when compression ratio was altered. To achieve this, the slope of the release changed: the higher the compression ratio, the steeper the slope when the release time was held constant. This interaction between release time and compression ratio might have had an effect that we were unable to measure or to account for in the analyses.

At present, the degree of change in release time needed to affect intelligibility significantly is not fully understood. The determination of such effects through future research could assist clinicians in achieving more successful hearing aid fittings. It is important to keep in mind, however, that the practice of setting a hearing aid to a prescriptive formula is a verification procedure (Kuk, 2001) and does not necessarily reflect the final fitting. The effects of release-time changes for a hearing aid programmed for more severe hearing losses and the determination of correction factors when setting slow-release-time hearing aids to nonlinear prescriptive targets are both areas that deserve further investigation.

REFERENCES


