Setting Hearing Aids Differently for Different Languages

Marshall Chasin, Au.D., Reg. CASLPO, Aud(C)[1]

ABSTRACT

In setting a hearing aid, the Speech Intelligibility Index (SII)—the percentage of speech cues that are audible—tells most of the story but is far from the entire picture. It can provide information on the various frequency importance bands in different languages (and thereby which phonemes or speech sounds are important), but it does not provide information on word-level and sentence-level cues that may be very important as well. Phoneme-level differences (seen in an SII) could result in a change in the frequency response and output specification for hearing aids such as an increase bass response in a language that has a larger number of sonorants than English or an increase in the frequency response in the 3000-Hz region for those (Slavic) languages that have palatalization as a distinctive cue. However, word- and sentence-level differences would not typically be seen in the SII measure and may necessitate changes in various compression-related parameters. For example, languages such as Japanese that have a rigid morphological consonant-vowel-consonant structure may necessitate a different release time for their nonlinear processing to maintain audibility of the quieter consonants. Sentence-level differences seen for subject-object-verb languages such as Hindi and Urdu may require increased amplification for soft-level inputs. Specific recommendations on how to program a hearing aid for many non-English languages are given, based on their phoneme-, word-, and sentence-level grammatical characteristics.

KEYWORDS: Languages, SVO, SOV, WDRC, release time, SII, hearing aids, program

Learning Outcomes: As a result of this activity, the participant will be able to (1) program a hearing aid differently depending on the language, (2) analyze a bilingual client's concerns about a poorly fit hearing aid, and (3) list differences between various commonly spoken languages

[1]Musicians' Clinics of Canada, Toronto, Ontario, Canada.

Address for correspondence and reprint requests: Marshall Chasin, Au.D., M.Sc., Reg. CASLPO, Aud(C), Doctor of Audiology, Director of Research, Musicians' Clinics of Canada, 340 College Street, Toronto, Ontario, Canada M5T 3A9 (e-mail: Marshall.Chasin@rogers.com).

Proceedings of the Widex Pediatric Audiology Congress; Guest Editor, André M. Marcoux, Ph.D.

Semin Hear 2011;32:182–188. Copyright © 2011 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New York, NY 10001, USA. Tel: +1(212) 584-4662.
DOI: http://dx.doi.org/10.1055/s-0031-1277240.
ISSN 0734-0451.

Of the more than 6800 languages spoken worldwide, only 4% of them are spoken by more than 95% of the people in the world. That is, the vast majority of languages are only spoken by several hundred or several thousand people. In contrast, only a few are spoken by more than 100 million people. Table 1 (data from the SIL Ethnologue Survey[1]) shows the distribution of the 11 most frequently spoken languages with approximate number of speakers as a first language. Many people are bilingual or even trilingual, and this list would be different if all of the speakers of a language were included. For example, English would be the most common language spoken because it is used as a language of commerce in many countries and because of the primarily English-based Internet.

Understandably, there may be some clinical concern about fitting hearing aids where the clients speak different languages. Byrne and his colleagues from around the world[2] studied the long-term average speech spectra (LTASS) of several different languages and found that "the similarity of the LTASS across samples demonstrates that it is reasonable to propose a universal LTASS which should be satisfactory for many purposes and applications to most, if not all, languages" (p. 2119). The finding of similar long-term speech spectra of different languages is understandable because speech emanates from a well-defined and almost identical vocal tract regardless of the language spoken and part of the world that any particular language is spoken. Adult humans have a vocal tract that is, on average, 17 cm in length between the vocal chords and the lips; a nasal cavity that is in parallel to the oral cavity; tongues; soft walled cheeks; and hard palates and soft palates. In short, as an acoustical device, the human vocal tract is very similar regardless of the speech that is generated by it and regardless of the location around the world.

The finding of Byrne at al[2] indicates that there should be no clinical concern about fitting hearing aids for different languages because the various long-term speech spectra appear to be similar for the various languages of the world. However, this provides no information about the various distributions of linguistically distinctive speech cues that are important for speakers of different languages. That is, there are clearly no significant phonetic differences between the various languages, but there may be significant phonemic differences. Fig. 1A shows the Speech Intelligibility Index (SII) of English,[3] and Fig. 1B shows the long-term average speech spectrum averaged over 10 languages.[2] For other languages, the SII would be different, despite having similar long-term speech spectra. The SII is the proportion of linguistically important speech cues that are audible in any one frequency band.[2] The SII is an updated version of an older American National Standards Institute (ANSI) standard called the Articulation Index, based in part on the work of Studebaker and Sherbecoe.[4]

Table 1 2009 Data from the Summer Institutes of Linguistics Ethnologue Survey

Rank	Language	Number of Speakers[Q1]
1	Chinese, Mandarin	1.12 billion
2	English	480 million
3	Spanish	320 million
4	Russian	285 million
5	French	265 million
6	Hindi-Urdu	250 million
7	Arabic	221 million
8	Portuguese	188 million
9	Bengali	185 million
10	Japanese	133 million
11	German	109 million

PHONEME-LEVEL DIFFERENCES—A FREQUENCY-RESPONSE CHANGE

Knowledge of the SII for any particular language is important and would provide valuable information regarding the shape of the aided response of the hard of hearing individual. Many hearing aid and real ear measurement manufacturers have, or are considering having, language-specific SII in their fitting and measurement software. However, the SII primarily only provides information about phoneme-level issues—an important first step. The SII is not sensitive to information concerning word-level or sentence-level issues that may

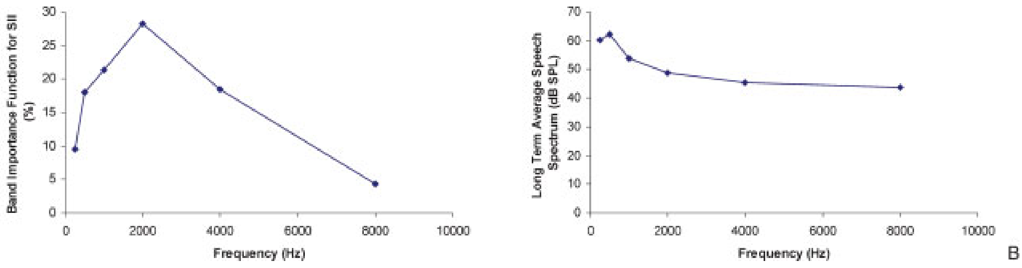

Figure 1 (A) Band importance function for the Speech Intelligibility Index (SII) of English. Adapted from American National Standards Institute.[3] (B) Long-term average speech spectrum averaged across 10 languages. Adapted from Byrne et al.[2]SPL[Q2].

Q2

affect the electroacoustic settings necessary for different languages. Strictly speaking, the SII as implemented by ANSI[3] is arrived at by examining a combination of band importance functions, word lists, as well as some conversational speech. The conversational speech aspect may contribute some linguistic information that is suprasegmental, or beyond the phoneme. Some research has appeared in the literature on languages such as Chinese.[5,6]

Other band-importance function implementations of future versions of the SII, or alternate calculations in different countries, may be different from the procedure set out in ANSI[3] and may not contain any information regarding linguistic elements beyond the level of the phoneme. The SII data shown in Fig. 1A are for English and according to the ANSI[3] procedure contain information regarding band importance, various word lists, and conversational speech.

Phoneme-level differences between various languages may result in differences in the shape of the frequency-response curve (and subsequently will have ramifications for the OSPL90[Q3] curve) for various languages. There will not typically be changes in the compression circuitry based on SII differences—these may be the result of word- and sentence-level differences between languages.

Q3

WORD-LEVEL DIFFERENCES—A RELEASE TIME COMPRESSION ISSUE

There are some differences between languages that may be seen, not just on the phoneme (SII) level but also at the level of the word, and even at the level of the sentence. Unlike phoneme-level changes, which can alter the shape of the frequency response, these suprasegmental differences may result in the selection of different settings for the compression circuitry.

There are several languages that have a highly constrained morphological structure. Unlike English where words may have two vowels or consonants in a row, languages such as Japanese and Vietnamese are constrained and must have a consonant-vowel-consonant (CVC) structure. In these languages, an intense vowel is followed by a less intense consonant (which in turn may be followed by an intense vowel). Even borrowed words into these languages must follow this rigid morphological structure. For example, the English word of the restaurant "MacDonald's" becomes "Macodonalodos" in Japanese with the occasional vowel [o] inserted to fulfill the morphological CVC requirement. This is also mostly true for Vietnamese, although there appears to be a syntactic and stylistic environment when this grammatical rule may be violated.

A hypothesis would therefore be that in Japanese and Vietnamese, for optimal audibility of the quieter (nonsonorant) consonants that follow a vowel, the release time of the compressor circuitry needs to be more rapid than for a typical English word. Preliminary data do support this hypothesis, but the current sample size is still too small to achieve statistical significance given the inherent variability in performing these clinical experiments. The ANSI requirements for the calculation of release time specifies a 50% tolerance and this adds to the variability.[7]

SENTENCE-LEVEL DIFFERENCES—GAIN FOR SOFT INPUTS ON COMPRESSION CIRCUITRY

Like the word-level differences (e.g., rigid CVC structure), language-specific sentence-level issues also can be of importance for setting some aspect of the compression circuitry. Specifically, for those languages that have a subject-object-verb word order, the amount of gain necessary for soft-level inputs on a wide dynamic range compression circuit needs to be increased relative to languages that have a subject-verb-object (SVO) word order such as English.

A typical English sentence may be "Catherine ate pizza." This has the structure of subject (Catherine) verb (ate) object (pizza). The vast majority of languages in the world are of this form and have sentence final objects. Objects, like all content nouns (subjects, pronouns, etc.) are more intense than function words such as verbs, prepositions, adverbs, adjectives, and conjunctions. As a result, the sentence final intensity is not as quiet as if the sentence would end in a non-noun. Sentence final intensity will always be quieter than sentence initial simply because the air flow from our lungs is decreasing as we run out of air. This is exacerbated when the sentence final item is a function word (non-noun) as found in some languages. The languages with a non-noun sentence final item are referred to as subject-object-verb (SOV) languages. Some examples of these languages are Japanese, Korean, Hindi-Urdu, and Turkish. A hallmark of most SOV languages is that they have postpositions rather than prepositions (with Somali being an exception). Our English sentence "Catherine ate pizza" would have the word order in Japanese of "Catherine pizza ate." If there was a preposition, such as "Catherine put the pizza in the oven" with *in* being the preposition, the Japanese word order would be "Catherine pizza oven in put." The "in put" portion of the sentence would be much less intense than the "in the oven" portion of the English word order sentence. Figure 2 shows a schematic of the intensity over time of an SVO sentence (with an English example in Fig. 3 and an SOV Korean example in Fig. 4).

The relevant hypothesis would be that people speaking and listening to SOV languages would require more gain for soft-level inputs (typically found in sentence final positions) than those speakers of SVO languages. As far as hearing aid technology is concerned, this may take various forms, from a reduction in the threshold kneepoint (TK), which would enhance the gain for low-level (sentence final) inputs as shown in Fig. 5, or, depending on the manufacturer, simply allow a greater gain specified for low-level inputs.

As a result of a pilot test on many hard-of-hearing people who spoke a wide range of SOV languages (e.g., Japanese, Korean, Hindi-Urdu, Turkish), it became apparent that there was no statistically significant difference between the

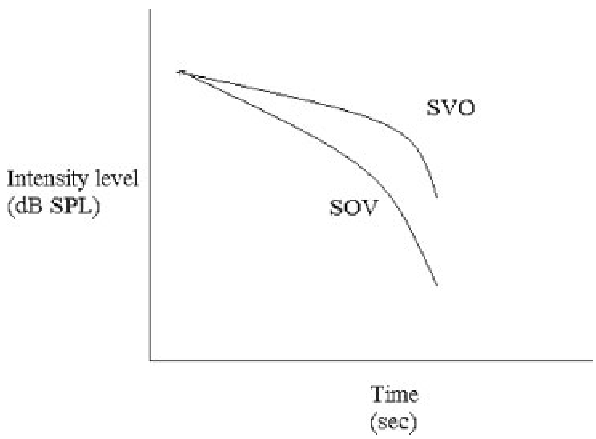

Figure 2 A stylized decrease in intensity for subject-verb-object (SVO) and subject-object-verb (SOV) languages as a function of time throughout the sentence. SPL[Q4].

Q4

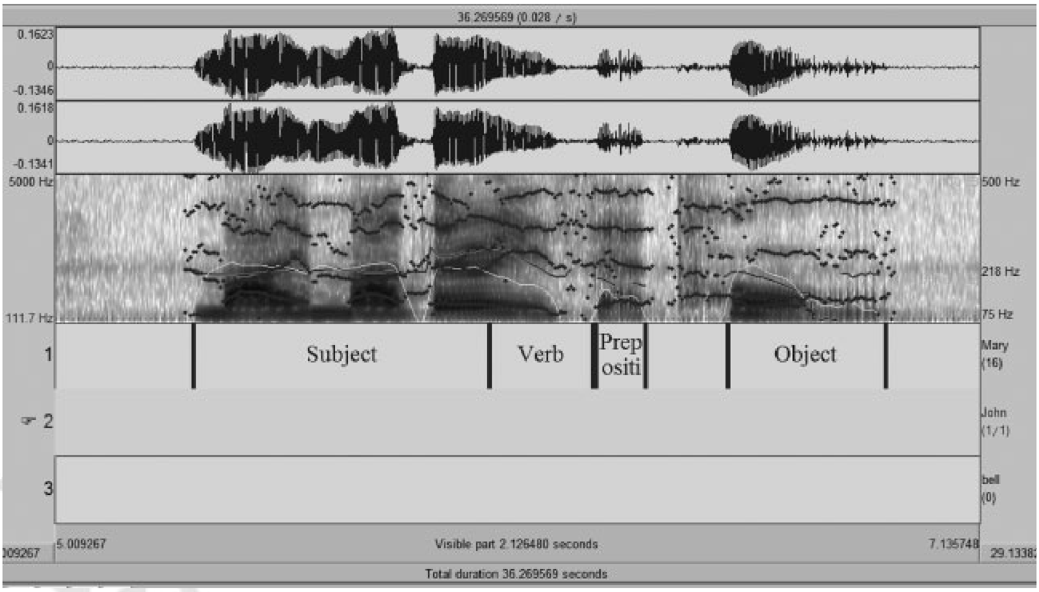

Figure 3 English: ''My mother is at home.'' Note the higher intensity of the sentence final object as compared with the sentence final elements in Figure 4. Unpublished data from Chasin (2010).

various SOV languages. They were merged and, as far as this study was concerned, were treated together as one language type. As part of the actual hearing aid fitting process in a clinical audiology facility, .wav files of various languages were played to bilingual hard-of-hearing people. Initially, program 1 was set according to the audiometric requirements of the person and their hearing loss for English. Once completed, clients were given control of the hearing aid fitting software module and allowed to set the TK or the equivalent method for the hearing aid software that would allow them to set the desired amount of gain for low-level inputs.

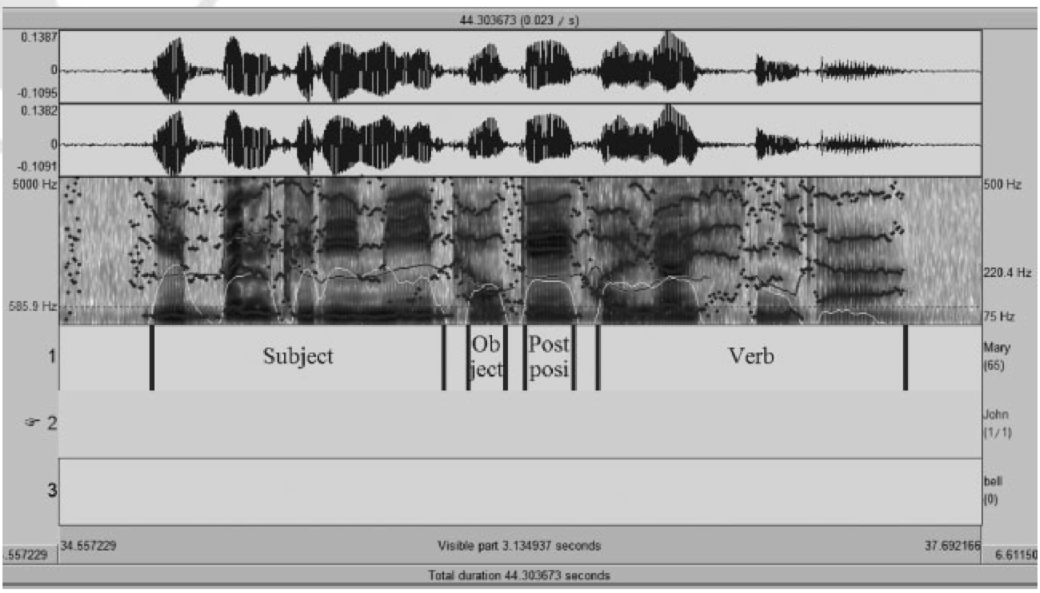

Figure 4 Korean: ''A pretty picture is hanging on the wall.'' Actual Korean word order: ''A pretty picture the wall on is hanging.'' The phrase ''on is hanging'' has significantly lower intensity than the sentence initial subject. Unpublished data from Chasin (2010).

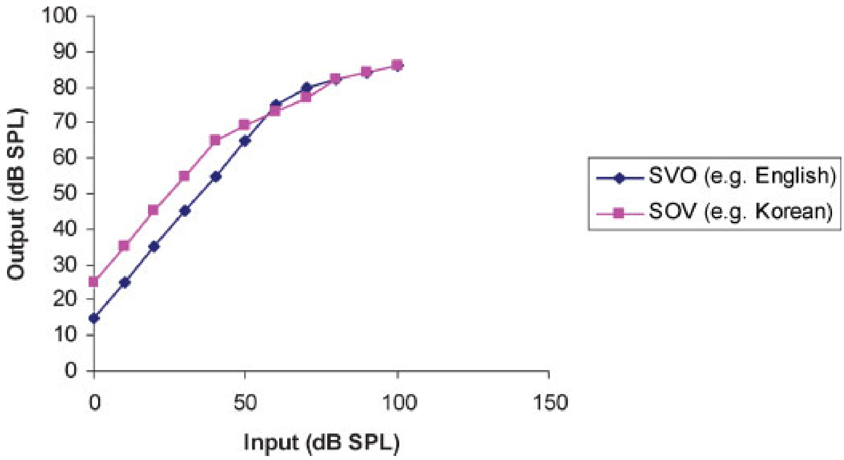

Figure 5 A lower-level threshold kneepoint setting for WDRC[Q5] on an input/output curve would result in more gain for softer inputs in a subject-object-verb (SOV) language such as Korean than for a subject-verb-object (SVO) language such as English. SPL[Q6]. Q5 Q6

The difference at 1000 Hz (or at 2000 Hz for those who required less than 15 dB of gain at 1000 Hz) between program 1 (set for English) and the amount of desired gain while listening to a .wav file of their second SOV language (program 2) was recorded. Using the subject as their own control has the advantage of minimizing error and controlling for factors that affects the specifying of gain and output for both programs, such as fitting formula. Statistically, this was analyzed by using a paired *t* test (and looking for differences within each individual).

To date, 86 people were studied (with only one ear being represented in those with a binaural hearing aid fitting). When listening to the SOV language, 5.4 dB more gain for low intensity speech was desired than when listening to English, and this was statistically significant at the $p < 0.05$ level (standard deviation 2.9 dB). That is, on average the "English" program was set with a higher TK (or equivalent) than the "Japanese" or other SOV language program for each person. The hard-of-hearing person can then use program 1 while listening to English and program 2 while listening to their second SOV language.

DISCUSSION

Certain general statements can be made about how the electroacoustic characteristics of hearing aids may be set for different languages. In most cases, program 1 may be set for English and program 2 may be set for a non-English language if there are significant differences in the SII of that second language or if there are other word-level or sentence-level grammatical requirements. The SII in its current implementation is only sensitive to phoneme-level differences and not to word- or sentence-level differences. Clearly more work needs to be performed on obtaining SII values for different languages. In the mean time, this summary, as well as the examples mentioned in Table 2 will provide the clinical audiologist with recommendations on how to set a hearing aid differently for a non-English language. Changes in the SII will result in changes in the shape of the frequency response, whereas changes in word- and sentence-level issues may result in a change in the compression circuitry behavior.

1. If a language has many nasal (e.g., Portuguese and Hindi-Urdu) or tonal (e.g., Chinese, Vietnamese, and Somali) consonants[Q9] or more timed[Q10] or vowel length importance (e.g., Japanese, Somali, and perhaps Spanish), then more gain is required than for English in the 125- to 2000-Hz region. This is also apparent on the SII. Q9 Q10
2. If retroflexion of the sonorant [r] is linguistically distinctive (e.g., Chinese), then more gain is required than for English in the 2700- to 3000-Hz region. This is also apparent on the SII.

Table 2 Examples of Languages That Possess Certain Non-English-Like Phonemic Patterns and Suggestions of How the Electroacoustic Settings May Be Altered Relative to English (This Will Be Apparent on Speech Intelligibility Index Measures)

Feature	Example	Difference from English
High-frequency consonants	Arabic	More gain > 2000 Hz
Q7 Nasal consonants[Q7]	Portuguese	More gain 125–2000 Hz
Palatal consonants	Russian "ch"	More gain 3000–3500 Hz
Retroflex consonants	Chinese "r"	More gain 2700–3000 Hz
Tonal consonants	Chinese	More gain 125–2000 Hz
Q8 Timed[Q8]	Japanese	More gain 125–2000 Hz

Adapted from Chasin.[8]

3. If palatal sounds are linguistically distinctive (e.g., Russian), then more gain is required than for English in the 3000- to 3500-Hz region. This is also apparent on the SII.
4. In languages that have a strict morphological CVCV structure (e.g., Japanese and Vietnamese), a shorter release time is required on the compression circuitry than for English. This is not apparent on the SII because it is a word-level issue.
5. In languages that have a SOV word order, typically with postpositions (e.g., Japanese, Korean, Hindi-Urdu, and Turkish), more gain is required for softer-level inputs on the Q11 compression (e.g., WDRC[Q11]) circuitry than for English. This is not apparent on the SII because it is a sentence-level issue.

REFERENCES

1. Summer Institutes of Linguistics (SIL) Ethnologue Survey. Summer Institutes of Linguistics 2009[Q12] Q12
2. Byrne D, Dillon H, Tran K, et al. An international comparison of long-term average speech spectra. J Acoust Soc Am 1995;96(4):2108–2120
3. American National Standards Institute (ANSI). ANSI S3.5, 1997—American National Standards for Calculation of the Articulation Index. New York, NY: ANSI; 1997
4. Studebaker GA, Sherbecoe RL. Frequency-importance and transfer functions for recorded CID W-22 word lists. J Speech Hear Res 1991;34(2):427–438
5. Kewley-Port D, Burkle TZ, Lee JH. Contribution of consonant versus vowel information to sentence intelligibility for young normal-hearing and elderly hearing-impaired listeners. J Acoust Soc Am 2007;122(4):2365–2375
6. Wong LL, Ho AH, Chua EW, Soli SD. Development of the Cantonese speech intelligibility index. J Acoust Soc Am 2007;121(4):2350–2361
7. American National Standards Institute (ANSI). ANSI S3.22–2003—Specification of Hearing Aid Characteristics. Melville, NY: Acoustical Society of America; 2003
8. Chasin M. How hearing aids may be set for different languages. Hearing Review 2008;15(11):16–20