Technical Addendum to the Familiar Sounds Audiogram (FSA)

Uchanski & Davidson; December 5, 2024

This document provides technical details and references to explain the rationale behind: 1) the shape, size and placement of the speech region; 2) the placement of the selected Ling 6 speech sounds; and 3) the placement of other familiar sounds along the outside of an audiogram. A few published studies offer background information on speech areas within audiograms, including their historical development (e.g., Tyler, 1979, as cited in Humes, 1991). Significant variations in how speech regions are depicted have been noted, largely due to differences in the chosen average speech level (Olsen, Hawkins & Van Tasell, 1987). More recent research highlights substantial inaccuracies and variability in familiar sounds audiograms currently in circulation, calling for revision and standardization (Hillis, Uchanski & Davidson, 2023).

1. **DESIGN OF SPEECH REGION ON AN AUDIOGRAM:** Depiction of a speech region on an audiogram requires knowledge of the mean spectrum of speech, the overall level of the speech, and the dynamic range of speech, as well as a conversion from dB SPL to dB HL.

    a. **Mean Spectrum of Speech, or Long-term Average Speech Spectrum (LTASS)**: The mean spectrum measurement shows the average intensity levels of speech, from a long recording, at different frequencies. The mean spectrum levels for this Familiar Sound Audiogram's (FSA's) speech region (which correspond to an approximate vertical mid-point of the speech region, prior to dB HL conversion) are *based on the Long-Term Average Speech Spectrum (LTASS) for adult female speech from the study of Byrne et al. (Table II, 1994)*. By design, this same female LTASS is used to create the ISTS test signal (Holube et al., 2010). Additionally, 1/3-octave band analysis shows that 90% of speech energy is concentrated in low frequencies, i.e., at or below 800 Hz.

    b. **Overall Level of the Speech:** *The overall level of conversational speech chosen for the speech region in this FSA is 60 dB SPL. This level corresponds to a typical vocal effort by the talker, and a distance of 1 m (~3-4 ft) between the talker and listener.* This overall level is supported by Boothroyd (2019), Olsen (1998), Pearsons, Bennett, & Fidell (1977), Cox, Matesich, & Moore (1988, p. 1101), Stelmachowicz, Lewis, Hoover, & Nishi (1993), Cornelisse, Gagné, & Seewald (1991, Table 1, p. 49), Holube, Fredelake, Vlaming, & Kollmeier (2010, p. 892), and Skinner, Holden, & Whitford (1997).

    c. **Dynamic Range of Speech:** *The dynamic range of speech, chosen for each 1/3-octave band in the speech region of this FSA, spans from +12 dB to -18 dB about the mean spectrum level.* This represents a dynamic range of 30 dB in each 1/3-octave band, which is consistent with the dynamic range of speech used in SII calculations and is supported by the Audioscan Verifit manual (2021, p. 96; +12 dB to -18 dB re: LTASS correspond to L1 [99th-percentile, aka speech peaks] and to L70 [30th-percentile, aka the valleys of speech]

respectively in each 1/3-octave band), Cox, Matesich, & Moore (1988, See Figure 1, p. 1101), Holube, Fredelake, Vlaming, & Kollmeier (2010, See Table 3, p.899 and Figure 9, p.900), and Byrne et al. (1994).

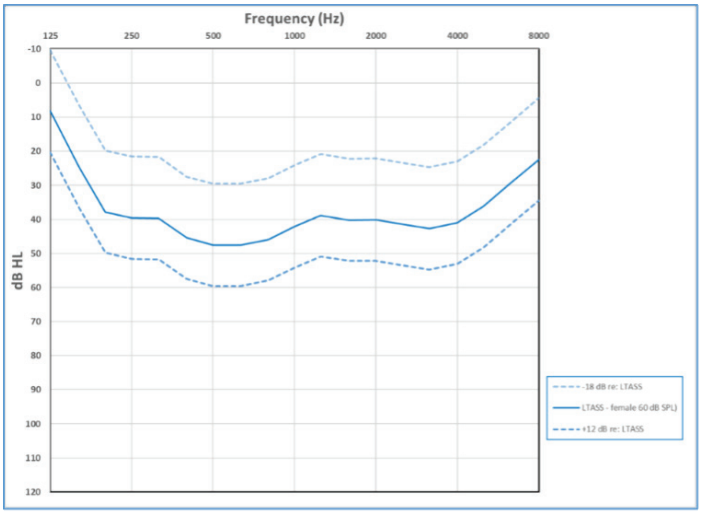
d. **Conversion from dB SPL to dB HL:** Thresholds for pure-tones are similar to thresholds for 1/3-octave-band- or octave-band-filtered signals. Justification for this statement is from the studies of Pascoe (1975, Figure 2), Abouchacra & Letowski (1999, Table 2), and Cox & McDaniel (1986, Tables I & II). Since thresholds for 1/3-octave band signals are similar to pure-tone thresholds, *minimum audible field (MAF) pure-tone thresholds, for binaural listening in the free field to signals at 0 degrees azimuth, are used to convert dB SPL to dB HL for 1/3-octave bands of speech.* Others (Pascoe, 1980; Popelka & Mason, 1987; Chial, 1998; Boothroyd, personal communication, March 2024) have used binaural MAF thresholds for this conversion. See Table 1, below, for a summary of the values for the speech region of this FSA. See also Figure 1 for a visual representation of the speech region plotted on an audiogram.

Table 1. Summary of values for the speech region in the FSA.

1/3-octave band center Frequency (Hz)	Mean Spectrum Level (dB SPL)[1]	RETSPL (dB re 20 µPa) for free field testing[2]	dB HL (LTASS, for avg 60 dB SPL)[3]	dB HL (weak edge of speech region)	dB HL (intense edge of speech region)
125	40.1	22.1	8	-10	20
160	53.4	17.9	25	7	37
200	62.2	14.4	38	20	50
250	60.9	11.4	40	22	52
315	58.1	8.6	40	22	52
400	61.7	6.2	45	27	57
500	61.7	4.4	48	30	60
630	60.4	3.0	48	30	60
800	58.0	2.2	46	28	58
1000	54.3	2.4	42	24	54
1250	52.3	3.5	39	21	51
1600	51.7	1.7	40	22	52
2000	48.8	-1.3	40	22	52
2500	47.3	-4.2	41	23	53
3150	46.7	-6.0	43	25	55
4000	45.3	-5.4	41	23	53
5000	44.6	-1.5	36	18	48
6300	45.2	6.0	29	11	41
8000	44.9	12.6	22	4	34

[1]Values are from Byrne et al (1994) for female speech with an overall level of 70 dB SPL.
[2]Reference equivalent threshold sound pressure levels (RETSPLs) (dB re 20 µPa) for free field testing from Table 9, ANSI S 3.6 (2010). These values are identical to those in the ISO 389-7 (2019) standard.
[3]Values are from the 2nd column, adjusted to an overall speech level of 60 dB SPL, then adjusted by the appropriate frequency-dependent RETSPL value, and rounded to the nearest integer.

Figure 1. Audiogram depicting the resultant dB HL values for the 60 dB female LTASS, with its dynamic range (+12 dB and -18 dB).



2. LING 6 SOUNDS:

a. FREQUENCY DATA:

i. **Vowel formant frequencies of adult female speech** are the average values from published works by Peterson & Barney (1952), Hillenbrand, Getty, Clark & Wheeler (1995), and Lee, Potamianos & Narayanan (1999) as shown in Table 2 below.

Table 2. Formant frequencies of adult female vowels /a/, /i/ and /u/ from three studies.

	Peterson & Barney (1952)		Hillenbrand et al. (1995)		Lee et al. (1999)		AVERAGE	
Vowel	F1 (Hz)	F2 (Hz)	F1 (Hz)	F2 (Hz)	F1 (Hz)	F2 (Hz)	F1 (Hz)	F2 (Hz)
/a/	850	1220	936	1551	894	1459	893	1410
/i/	310	2790	437	2761	360	2757	369	2769
/u/	370	950	459	1105	412	1388	414	1148

ii. **Nasal /m/ of adult female speech:** The spectral peak of /m/ is based on the published works by Pittman, Stelmachowicz, Lewis & Hoover (2003, Fig. 3, p. 653), Boothroyd, Erickson & Medwetsky (1994, Table 1, p. 436), and Scollie et al. (2012). The primary spectral peak for /m/ is 250 Hz in all three studies.

iii. **Fricatives of adult female speech:** Spectral peaks for fricatives are the average values from the published works by Pittman et al. (2003, p.654), Jongman, Wayland & Wong (2000, Figure 2), and Scollie et al. (2012, Figure 1). Values are shown in Table 3.

Table 3. Peak frequencies of /sh/ and /s/ for adult female speech from three studies.

	Pittman et al. (2023)	Jongman et al. (2000)	Scollie et al. (2012)	AVERAGE
Fricative	Peak (Hz)	Peak (Hz)	Peak (Hz)	Peak (Hz)
/sh/	4500	4332	2500	3777
/s/	7300	7496	6360	7052

b. **LEVEL DATA:** The *absolute* levels and the *relative* levels (of vowels to other vowels, of a vowel's F1 to that vowel's F2, and of consonants to vowels) for the Ling 6 sounds on this FSA were determined using the assumptions and steps described below.

i. The overall level of speech, for the speech region, is 60 dB SPL.

ii. For typical vocal effort speech, vowels usually fall within a few dB of each other in their overall level (See House & Fairbanks, 1953, Fig. 4; Koenig & Fuchs, 2019, Fig. 1; Horii, House & Hughes, 1971, Fig. 5). Thus, relative to each other, all three Ling vowels will have the same overall level.

iii. The levels of the three vowels are also assumed to be the same as the overall speech level of 60 dB SPL.

iv. The *relative* levels of a vowel's formants (level of F2 relative to the level of F1) are based on the levels of vowel formant spectral peaks in 1/3-octave spectra reported by Scollie et al. (2012), and on the A1 and A2 values from Peterson & Barney (1952, Table II, p. 183). See Table 4 below.

Table 4. Levels of Vowel Formants for /a/, /i/, and /u/, and Average Relative Levels across Two Studies. Amplitude data from Peterson & Barney are relative to the level of the first formant amplitude of the /ɔ/ vowel for each speaker. Data from Scollie et al. (2012) are after an adjustment to an overall vowel level of 60 dB SPL.

Vowel	Level of First Formant (F1)		Level of Second Formant (F2)		Level of F2 to *relative* to Level of F1
	Peterson & Barney (1952); (dB)	Scollie et al. (2012); (dB SPL)	Peterson & Barney (1952); (dB)	Scollie et al. (2012); (dB SPL)	AVERAGE (dB)
/a/	-1	54	-5	54	-2
/i/	-4	56	-24	52	-12
/u/	-3	55	-19	41	-15

v. The *relative* level of F1 to the overall vowel level is estimated from data of Scollie et al (2012) and of Peterson & Barney (1952). In Scollie et al., these three vowels' F1 levels are 4 to 6 dB below the overall vowel level, while in Peterson & Barney, F1 levels are 1 to 4 dB below the first formant amplitude of the /ɔ/ vowel, the most intense vowel in spoken English. In this FSA, *F1 levels are set at 3 dB below the overall vowel level, for all three vowels.*

vi. The *relative* levels of the Ling consonants to the overall vowel level are based on data from Levitt (1978, Chapter 3, Table 2, p. 96), Gordon-Salant (1986), and Pittman et al. (2003, Figures 3-5) for /m/, and additionally on data from Jongman et al (2000) and Freyman & Nerbonne (1989) for the fricatives. Average consonant-to-vowel ratios (CVRs) from these studies are:

- /m/: -5.9 dB
- /sh/: -7.9 dB
- /s/: -14.9 dB

vii. The 1/3-octave band spectra of the Ling 6 sounds from Scollie et al. (2012) were used to estimate dB SPL levels in specific spectral regions, which were then converted to dB HL. See Table 5 below.

1. For the three vowels, their 1/3-octave band spectra were adjusted to the assumed overall vowel level of 60 dB SPL. The dB SPL levels in the first formant frequency region were then reduced by 3 dB (See step 2.b.v above), and then the levels in the 2nd formant frequency region were reduced, in dB, by values shown in Table 4. Then, the F1 and F2 levels in dB SPL were converted to dB HL using the ANSI thresholds (RETSPLs in Table 1) nearest in frequency to the corresponding 1/3-octave band center frequency and rounded to the nearest integer dB value.

2. For each of the three Ling consonant sounds, their 1/3-octave band spectra were adjusted to the average vowel level (60 dB SPL), and then additionally adjusted by the appropriate consonant-to-vowel ratio (CVR) (e.g., for /m/, after adjustment to 60 dB SPL, then subtract 5.9 dB). Then, the level in dB SPL at that consonant's spectral peak was converted to dB HL, using the ANSI threshold (RETSPL in Table 1) nearest in frequency to the corresponding 1/3-octave band center frequency and rounded to the nearest integer dB value.

Table 5. Frequencies and Levels of Ling 6 Sounds on this FSA. Note: The Ling Sounds listed phonetically in the table below correspond to their respective pronunciations on the FSA: /a/ as "ah", /i/ as "ee", /u/ as "oo", /m/ as "m", /sh/ as "sh", and /s/ as "s".

Ling sound	F1 (Hz)	F1 level (dB SPL)	F1 level (dB HL)	F2 (Hz)	F2 level (dB SPL)	F2 level (dB HL)	Spectral peak (Hz)	Level of spectral peak (dB SPL)	Level of spectral peak (dB HL)
/a/	893	51	49	1410	49	46			
/i/	369	53	41	2769	41	45			
/u/	414	55	49	1148	38	35			
/m/							250	52	40
/sh/							3777	48	52
/s/							7052	44	38

3. POSITION OF TWO FAMILIAR SOUNDS ALONG THE OUTSIDE OF THE AUDIOGRAM.

Images and labels of two familiar sounds appear to the left of the y-axis of the FSA, namely "whisper" (Peterson & Gross, 1972) and "gas lawn mower" (Berger & Neitzel, 2016). These are intended to serve as anchors in sound level; they convey well-known 'soft' and 'loud' sounds at 'higher up' and 'lower down' positions along the audiogram dB HL axis. These are positioned to the left of the y axis partly to minimize visual clutter within the audiogram and partly due to the broad spectral energy of each of these two sounds. The vertical placement of these two familiar sounds, along the y-axis of the FSA, required knowledge of:

a. **Distance from sound source for the reported sound level** - Because the speech region of the FSA represents conversational speech levels at a distance of ~ 3ft, the sound levels of these two familiar sounds are also presented at a roughly comparable distance of 1 m.

b. **dB SPL to dB HL Conversion** – The difference between dB SPL (Sound Pressure Level) and dB HL (Hearing Level) varies across frequencies. For the purposes of this Familiar Sounds Audiogram (FSA), we applied an adjustment of 15 dB (dB SPL – 15 = dB HL), used for speech (Killion & Mueller, 2010), to both familiar sounds. No adjustment was made for the A-weighting of the dB SPL measurements of either sound.

   i. **Whisper:** Peterson & Gross (1972) measured the whisper at a distance of 150 cm. Using the 'inverse square law' for sound levels, which is a reasonable approximation for sound propagation in large unimpeded spaces, the 30 dBA sound level at 150 cm would be equated to a 33.5 dBA sound level at a distance of 1 m (30 dBA + 20*log(1.5 m/1 m). The 33.5 dBA measurement was then converted to 18.5 dB HL by applying the 15 dB adjustment (dB SPL – 15 = dB HL).

   ii. **Gas Lawn Mower:** Berger & Neitzel (2016) measured the level of the gas lawn mower at a distance of 1 m; hence, no adjustment for distance was necessary. The recorded level of 99 dBA was converted to 84 dB HL by applying the 15 dB adjustment (dB SPL – 15 = dB HL).

REFERENCES:

Abouchacra, K. S., & Letowski, T. (1999). Comparison of air-conduction and bone-conduction hearing thresholds for pure tones and octave-band filtered sound effects. *J Am Acad Audiol*, 10(8), 422-428.

ANSI S3. 6-2010. (2010). Specification for audiometers.

Berger, E.K.C., & Neitzel, R.L. "Noise Navigator Sound Level Database." Published 22 August 2016. Downloaded from https://multimedia.3m.com/mws/media/8885530/noise-navigator-sound-level-hearing-protection-database.pdf website on 14 November 2024.

Boothroyd, A. (2019). The Acoustic Speech Signal. In J. R. Madell, C. Flexer, J. Wolfe, & E. C. Schafer (Eds.), *Pediatric Audiology: Diagnosis, Technology, and Management* (3rd ed., pp. 207-213). Thieme. ISBN: 978-1-62623-402-4

Boothroyd, A., Erickson, F. N., & Medwetsky, L. (1994). The Hearing-Aid Input - a Phonemic Approach to Assessing the Spectral Distribution of Speech. *Ear and Hearing*, 15(6), 432-442. https://doi.org/Doi 10.1097/00003446-199412000-00004

Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R., Hagerman, B., Hetu, R., Kei, J., Lui, C., Kiessling, J., Kotby, M. N., Nasser, N. H. A., El Kholy, W. A. H., Nakanishi, Y., Oyer, H., Powell, R., Stephens, D., Meredith, R., . . . Ludvigsen, C. (1994). An international comparison of long-term average speech spectra. *Journal of the Acoustical Society of America*, 96(4), 2108-2120.

Chial, M. R. (1998). Yet Another Audiogram. *ASHA Special Interest Division 6: Hearing and Hearing Disorders: Research and Diagnostics*, 2-3.

Cornelisse, L. E., Gagné, J.-P., & Seewald, R. C. (1991). Ear Level Recordings of the Long-Term Average Spectrum of Speech. *Ear & Hearing*, 12(1), 47-54.

Cox, R. M., & McDaniel, D. M. (1986). Reference equivalent threshold levels for pure tones and 1/3-oct noise bands: insert earphone and TDH-49 earphone. *Journal of the Acoustical Society of America*, 79(2), 443-446.

Cox, R. M., Matesich, J. S., & Moore, J. N. (1988). Distribution of Short-Term Rms Levels in Conversational Speech. *Journal of the Acoustical Society of America*, 84(3), 1100-1104. https://doi.org/Doi 10.1121/1.396697

Freyman, R.L., & Nerbonne, G.P. (1989). The importance of consonant-vowel intensity ratio in the intelligibility of voiceless consonants. *J Speech Hear Res*, 32(3), 524-35. Doi: 10.1044/jshr.3203.524.

Gordon-Salant, S. (1986). Recognition of natural and time/intensity altered CVs by young and elderly subjects with normal hearing. *Journal of the Acoustical Society of America*, 80(6), 1599-1607.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, 97(5, Pt. 1), 3099-3111.

Hillis, C. L., Uchanski, R. M., & Davidson, L. S. (2023). Common Sounds Audiograms: Quantitative Analyses and Recommendations. *Semin Hear*, 44(Suppl 1), S49-S63. https://doi.org/10.1055/s-0043-1764128

Hirsh, I. J., Reynolds, E. G., & Joseph, M. (1954). Intelligibility of Different Speech Materials. *Journal of the Acoustical Society of America*, 26(4), 530-538.

Holube, I., Fredelake, S., Vlaming, M., & Kollmeier, B. (2010). Development and analysis of an International Speech Test Signal (ISTS). *International Journal of Audiology*, 49(12), 891-903. https://doi.org/10.3109/14992027.2010.506889

Horii, Y., House, A. S., & Hughes, G. W. (1971). A masking noise with speech-envelope characteristics for studying intelligibility. *J Acoust Soc Am*, 49(6), 1849-1856.

House, A. S., & Fairbanks, G. (1953). The Influence of Consonant Environment Upon the Secondary Acoustical Characteristics of Vowels. *Journal of the Acoustical Society of America*, 25(1), 105-113. https://doi.org/Doi 10.1121/1.1906982

Humes, L. E. (1991). Understanding the speech-understanding problems of the hearing impaired. *J Am Acad Audiol*, 2(2), 59-69.

ISO 389-7. (2019). Acoustics - Reference zero for the calibration of audiometric equipment. Part 7: Reference threshold of hearing under free-field and diffuse-field listening conditions. *International Organization for Standardization.*

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America*, 108(3), 1252-1263.

Killion, M. C., & Mueller, H. G. (2010). Twenty years later: A new count-the-dots method. *The Hearing Journal*, 63(1), 10-12.

Koenig, L. L., & Fuchs, S. (2019). Vowel Formants in Normal and Loud Speech. *J Speech Lang Hear Res*, 62(5), 1278-1295. https://doi.org/10.1044/2018_JSLHR-S-18-0043

Lee, S., Potamianos, A., & Narayanan, S. (1999). Acoustics of children's speech: Developmental changes in temporal and spectral parameters. *Journal of the Acoustical Society of America*, 105(3), 1455-1468.

Levitt, H. (1978). The Acoustics of Speech Production. In M. Ross & T. G. Giolas (Eds.), *Auditory Management of Hearing-Impaired Children* (pp. 45-115). University Park Press.

Olsen, W. O. (1998). Average Speech Levels and Spectra in Various Speaking/Listening Conditions. *American Journal of Audiology*, 7(October), 21-25. https://doi.org/10.1044/1059-0889

Olsen, W. O., Hawkins, D. B., & Van Tasell, D. J. (1987). Representations of the long-term spectra of speech. *Ear Hear*, 8(5 Suppl), 100S-108S.

Pascoe, D. P. (1975). Frequency responses of hearing aids and their effects on the speech perception of hearing-impaired subjects [Research Support, U.S. Gov't, P.H.S.]. *Ann Otol Rhinol Laryngol*, 84(5 pt 2 Suppl 23), 1-40.

Pascoe, D. P. (1978). An approach to hearing aid selection. *Hearing Instruments*, 29, 12-16.

Pascoe, D. P. (1980). Clinical Implications of Nonverbal Methods of Hearing Aid Selection and Fitting. *Seminars in Speech, Language and Hearing*, 1(3), 217-228.

Pearsons, K. S., Bennett, R. L., & Fidell, S. (1977). *Speech Levels in Various Noise Environments.*

Peterson, G. E., & Barney, H. L. (1952). Control Methods Used in a Study of the Vowels. *Journal of the Acoustical Society of America*, 24(2), 175-184.

Peterson, A. P., & Gross, E.E. Jr. (1972). *Handbook of noise measurement* (7th ed.). General Radio Company.

Pittman, A. L., Stelmachowicz, P. G., Lewis, D. E., & Hoover, B. M. (2003). Spectral Characteristics of Speech at the Ear: Implications for Amplification in Children. *Journal of Speech, Language, and Hearing Research*, 46(June), 649-657.

Popelka, G. R., & Mason, D. I. (1987). Factors which affect measures of speech audibility with hearing aids. *Ear Hear,* 8(5 Suppl), 109S-118S.

Scollie, S., Glista, D., Tenhaaf, J., Dunn, A., Malandrino, A., Keene, K., & Folkeard, P. (2012). Stimuli and normative data for detection of Ling-6 sounds in hearing level [Research Support, Non-U.S. Gov't]. *Am J Audiol*, 21(2), 232-241. https://doi.org/10.1044/1059-0889(2012/12-0020)

Skinner, M. W., Holden, L. K., Holden, T. A., Demorest, M. A., & Fourakis, M. S. (1997). Speech recognition at simulated soft, conversational and raised-to-loud vocal efforts by adults with cochlear implants. *Journal of the Acoustical Society of America*, 101(6), 3766-3782.

Stelmachowicz, P. G., Mace, A. L., Kopun, J. G., & Carney, E. (1993). Long-Term and Short-Term Characteristics of Speech: Implications for Hearing Aid Selection for Young Children. *Journal of Speech and Hearing Research*, 36(June), 609-620.

Tyler, R. S. (1979). Measuring Hearing Loss in The Future. *British Journal of Audiology*, 13, 29-40.