Potential Mechanisms for Perception of Frequency-Lowered Speech

Joshua M. Alexander
Ph.D., CCC-A

PURDUE UNIVERSITY

Disclosures

- I have received past research support from
  - The National Institutes of Health
  - The Indiana Clinical and Translational Sciences Institute
  - The William Demant Holding Group
- I hold a patent on a novel form of frequency lowering
- I am a paid consultant for Creare, LLC on an NIH SBIR project to develop an open-source, master hearing aid
- I have received speaker honoraria from Signia, Oticon, Phonak, Starkey, and ReSound hearing aid companies

2

Acknowledgements

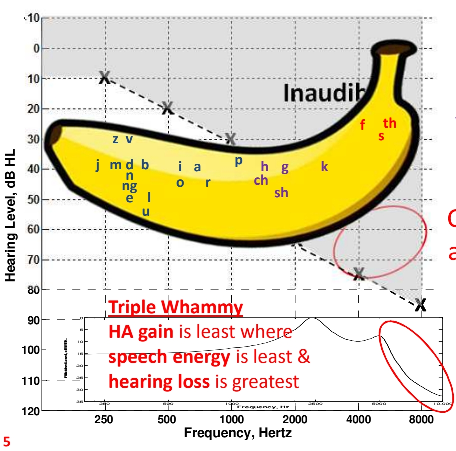
3

Agenda

- **Key point:** *frequency-lowering methods and settings that preserve the greatest amount of temporal modulation from the original speech at the level of the auditory periphery will yield the best outcomes for speech perception*

1. **Big picture**: who, what, and why of frequency lowering
2. **Paradigm shift** from the frequency domain to the time domain for developing a theoretical model that describes perception of frequency-lowered speech
3. Evaluate different **models/mechanisms** using the newest method of frequency lowering
4. Critically evaluate current **clinical recommendations** for selecting frequency lowering parameters

4

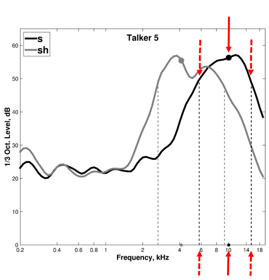


Add **gain** for audibility

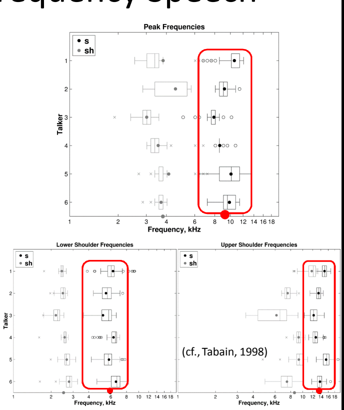
Conventional amplification is limited

5

Ultra High-Frequency Speech



Alexander (*in revision*)
"The S-SH Confusion Test and the Effects of Frequency Lowering"



6

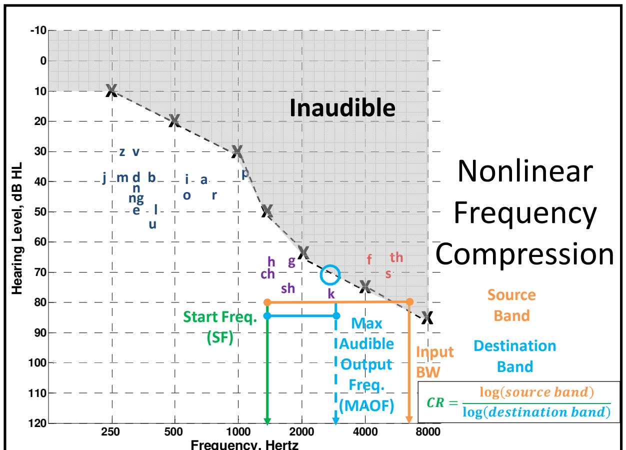
Frequency Lowering is Ubiquitous

(Alexander, 2016b)

Manufacturer	Feature Name	Frequency Lowering Method
Widex	Audibility Extender	Transposition (static)
	Enhanced Audibility Extender	Transposition (adaptive)
Phonak	SoundRecover [a]	Compression (static)
	SoundRecover2	Compression (adaptive)
Starkey	Spectral iQ [b]	Spectral Envelope Warping
Signia	Frequency Compression [c]	Compression
ReSound	Sound Shaper [d]	Proportional Compression
Oticon	Speech Rescue [e]	Multilayered Transposition

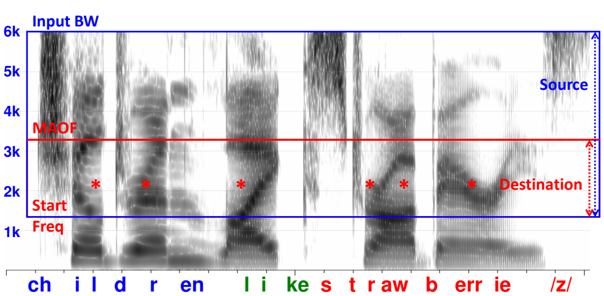
[a] Also offered by Unitron as "Frequency Compression" and by Hansaton as "Sound Restore"
[b] Also offered by Microtech as "Sound Compression"
[c] Also offered by Rextron as "Bandwidth Compression"
[d] Also offered by Beltone as "Sound Shifter" and by Interton as "Frequency Shifter"
[e] Similar to, but not the same as, what is offered by Bernafon as "Frequency Composition" and by Sonic as "Frequency Transfer"

7



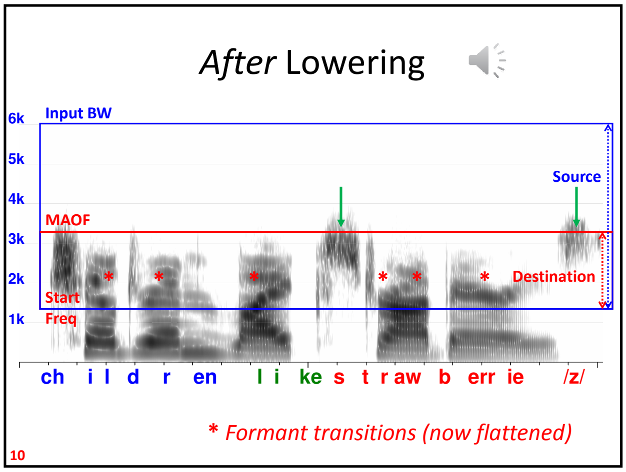
8

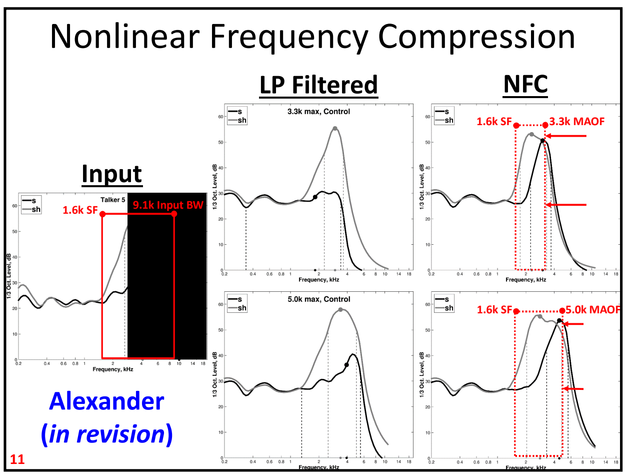
Before Lowering

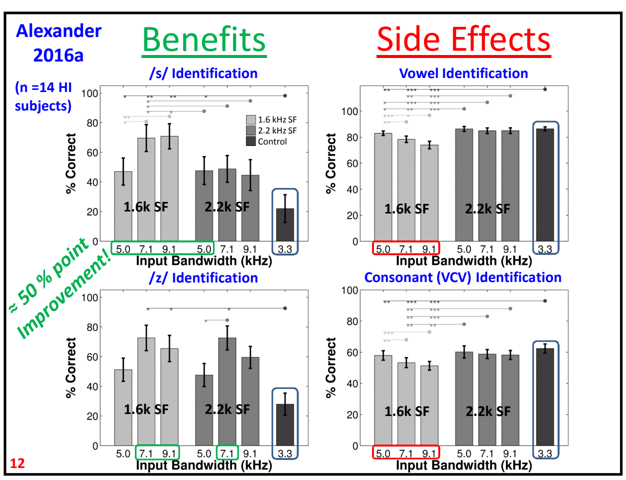


* *Formant transitions*

9

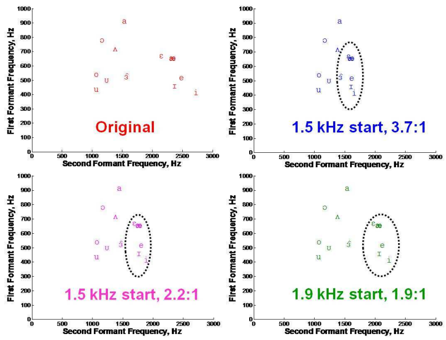

After Lowering
6k
5k
4k
3k
2k
1k
Input BW
Source
MAOF
Start
Freq
Destination
ch i l d r en l i ke s t r aw b err ie /z/
* Formant transitions (now flattened)
10


Nonlinear Frequency Compression
Input
LP Filtered
NFC
1.6k SF
9.1k Input BW
3.3k max, Control
5.0k max, Control
1.6k SF
3.3k MAOF
1.6k SF
5.0k MAOF
Alexander
(*in revision*)
11


Alexander
2016a
(n =14 HI
subjects)
Benefits
Side Effects
/s/ Identification
Vowel Identification
/z/ Identification
Consonant (VCV) Identification
≈ 50 % point
Improvement!
1.6 kHz SF
2.2 kHz SF
Control
% Correct
1.6k SF
2.2k SF
Input Bandwidth (kHz)
12

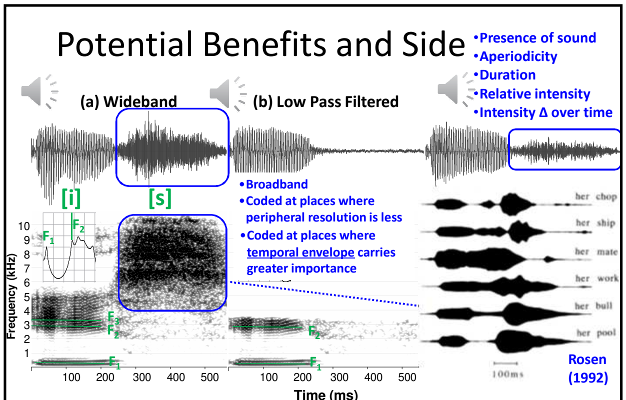
Potential Side Effects

Formant alteration, vowel reduction with a low cut off (**start)**



13

Potential Benefits and Side



(Alexander, 2016a)

14

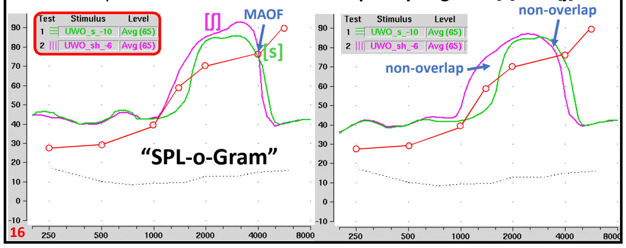
Paradigm Shift

- **From frequency domain to time domain**
  - Some transposition and <u>all compression methods</u> adhere to an **untested assumption** that <u>spectral features</u> of high-frequency speech cues need to be preserved or replicated after lowering
  - Previous attempts to 'optimize' in terms of frequency
    - Input bandwidth (*i.e.*, upper freq. limit)
    - Frequency separation (*i.e.*, between key speech contrasts, namely [s] and [ʃ])
- Preservation of features in the temporal modulation is ***not*** **a new concept** with respect to freq. lowering
  - It was the very premise behind most early methods of frequency transposition, especially <u>vocoding</u> methods

15

Recommended Clinical Procedure

- Scollie recommends a 2-step procedure using calibrated [s] and [ʃ] (e.g., Scollie *et al.*, 2016)
  1. Make high-frequency 'edge' of [s] audible (≤ MAOF)
  2. Make high-frequency edge of [s] and [ʃ] non-overlapping, if possible
     - Phonak also recommends at least 1/3-octave (≈1 ERB) separation between the **low-frequency edges** of [s] and [ʃ]



16

Untested Assumptions

1. [s]-[ʃ] discrimination is the same on the low-frequency edge as the high-frequency edge
2. [s]-[ʃ] edge frequency discrimination is constant across frequency and level
3. Maximizing [s]-[ʃ] edge frequency differences will maximize their discriminability
4. Maximizing [s]-[ʃ] discriminability will maximize perception of other fricatives and consonants

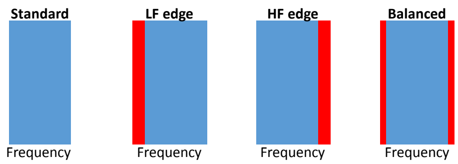
17

Unanswered Questions

1. What is the minimum frequency difference (**just noticeable difference, JND**) needed for normal-hearing listeners to tell two bands of noise apart?
2. How about for hearing-impaired listeners?
3. How does the addition of intervening sounds influence the JND for noise discrimination

18

Bandwidth Discrimination

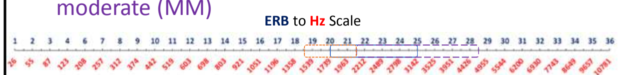


- 12 normal-hearing (NH), college-aged listeners
- Filtered pink noise samples (level/ERB ≈ constant)
  - **2 levels: 55 and 75 dB SLP/ERB**
  - 150 msec (25-msec ramps) ≈ duration of medial [s], [ʃ]
- 3-interval, forced choice task
  - For each trial, the noise samples were **frozen**
  - Listener had to identify the interval with the 'oddball'
    - (a) wider LF edge, (b) wider HF edge, (c) both edges balanced

19

Bandwidth Discrimination

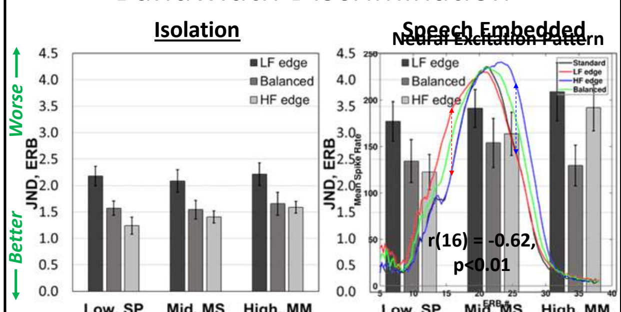
- 3 frequency ranges for the standard
  - Roughly correspond to **average frequency and BW of lowered [s],[ʃ]** for frequency-lowering settings appropriate for hearing losses that are (1) severely-profound (SP), (2) moderately-severe (MS), (3) mild-to-moderate (MM)



  - +/- 1 ERB rove on each trial (forced *within* trial comparisons)
  - 9 interleaved tracks: 3 freq ranges x {LF edge, HF edge, balanced} → tough!
- 2 down, 1 up steps (≈70.7% correct)
  - $\frac{1}{2}$ ERB (rev. 1-4), $\frac{1}{4}$ ERB (rev. 5-8), $\frac{1}{8}$ **ERB (rev. 9-16)**

20

Bandwidth Discrimination



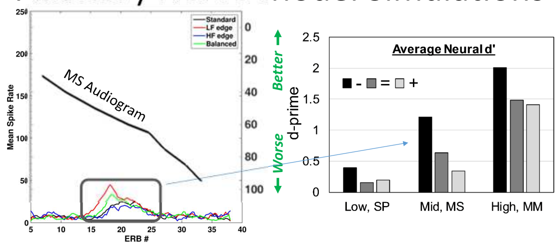
* No main effect for level

* No main effect for [low, mid, high]

* Main effect for edge freq
LF > Balanced > HF

21

Auditory Nerve Model Simulations

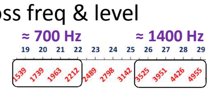


- AN model simulations strongly suggest that HI subjects will have an <u>opposite pattern</u> of results as NH subjects
  - Indicates that discrimination of frequency-lowered frication using **spectral separation alone will be difficult**, especially on the HF edge, and as hearing loss increases

22

Bandwidth Discrimination Summary

- **Tested Assumptions**
  1. Discrimination is the same on the low-frequency edge as the high-frequency edge
     - **FALSE**, discrimination is better on the high-freq edge
  2. Discrimination is constant across freq & level
     - **True** (for noise in isolation)



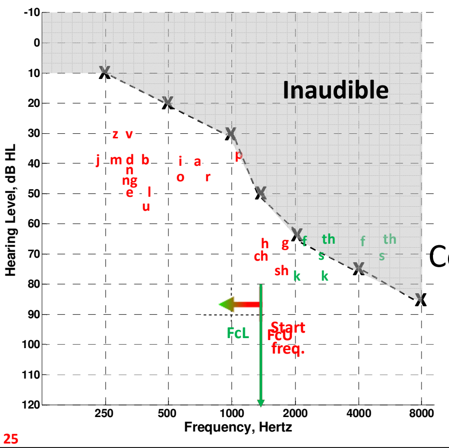
- **Answered Questions**
  1. Minimum freq. difference (JND) needed for normal-hearing listeners to tell two bands of noise apart?
     - ≈1.4 ERB (HF edge), ≈1.6 ERB (Balanced), ≈2.15 ERB (LF edge)
  2. *How about for hearing-impaired listeners?*
     - ***??? <u>Cannot</u> use HF edge!***
  3. How does the addition of intervening sounds influence the JND for noise discrimination
     - ↑ by ≈1.25 ERB, some effects of level and freq

23

Untested Assumptions

1. [s]-[ʃ] discrimination is the same on the low-frequency edge as the high-frequency edge
2. [s]-[ʃ] edge frequency discrimination is constant across frequency and level
3. Maximizing [s]-[ʃ] edge frequency differences will maximize their discriminability
4. Maximizing [s]-[ʃ] discriminability will maximize perception of other fricatives and consonants

24

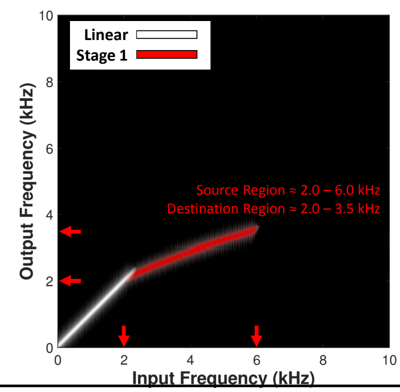


Adaptive

Nonlinear Frequency Compression

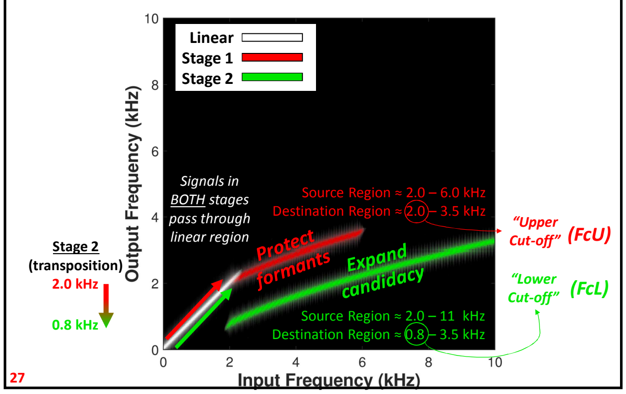
25

Conventional NFC



26

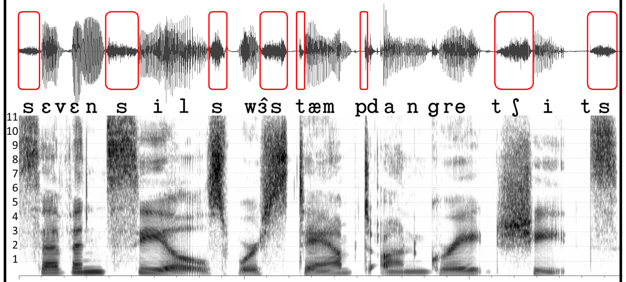
Adaptive NFC (ANFC)



27

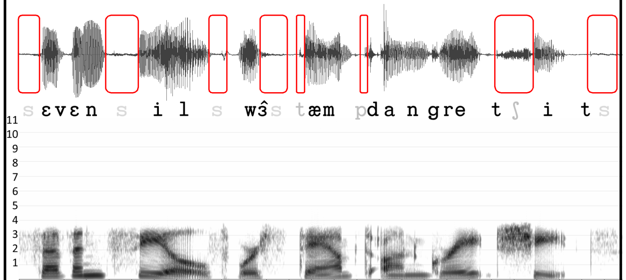
Full Bandwidth

"seven seals were stamped on great sheets"



28

Simulated Loss (Filtered >3.5kHz)



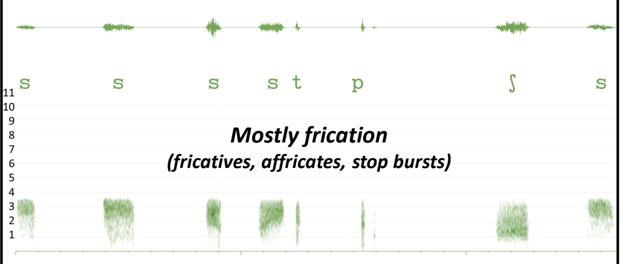
29

Processed with ANFC (Stage 1)

Stage 1
Source Region ≈ 2.0 – 6.0 kHz
Destination Region ≈ 2.0 – 3.5 kHz

30

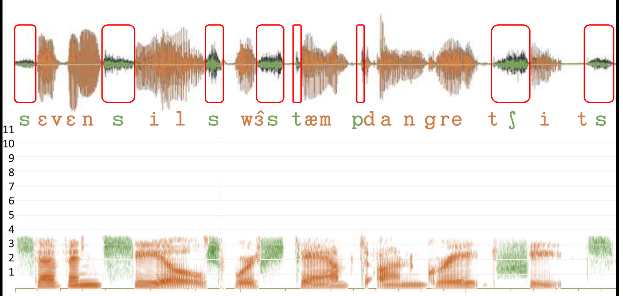
Processed with ANFC (Stage 2)



Stage 2
Source Region ≈ 2.0 – 11 kHz
Destination Region ≈ 0.8 – 3.5 kHz

31

Processed with ANFC



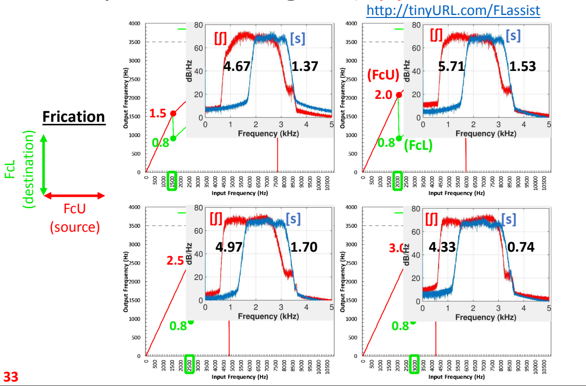
Stage 1
Source Region ≈ 2.0 – 6.0 kHz
Destination Region ≈ 2.0 – 3.5 kHz

Stage 2
Source Region ≈ 2.0 – 11 kHz
Destination Region ≈ 0.8 – 3.5 kHz

32

Source Region for Stage 2 (lower cutoff) *Depends* on Stage 1 (upper cutoff)

http://tinyURL.com/FLassist

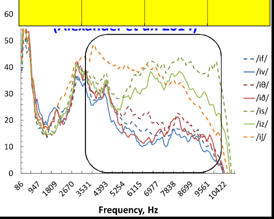


33

Test Stim

1. [s]-[ʃ] discrimination in words
   - s-sh Confusion Test (Alexander,
   - 66 word pairs (132 total), 1 woman ta
2. Fricative discrimination (/iC/)
   - 7 fricatives, 3 women talkers
3. Consonant discrimination
   - 20 consonants, 3 vowels, 1 m
4. Vowel discrimination (/hV
   - 12 vowels, 2 men + 2 women

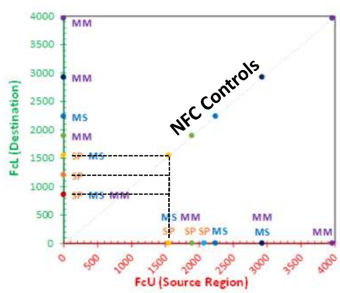
had (hat)	hod (hot)	hawed (taught)	head (bed)
hayed (hate)	heard (hurt)	hid (hit)	heed (heat)
hoed (road)	hood (hook)	hud (hut)	who`d (hoot)



34

Experimental Conditions

- 3 groups of NH listeners (12 each)
  - Bandwidth limited to 2.1 kHz (SP), 3.3 kHz (MS), or 5.0 kHz (MM)
- 8-9 different conditions processed in MATLAB that varied lower cut-off (FcL) and upper cut-off (FcU) with identical output bandwidth

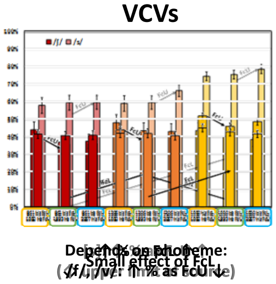


35

Results

Increasing the max input beyond what is minimally necessary → blurring of the signal in time and frequency

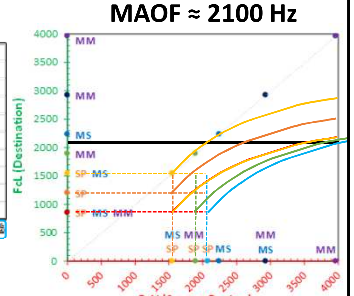
VCVs



Depends on phoneme:

Small effect of FcL

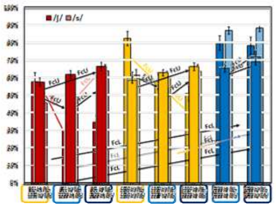
Severe-to-Profound MAOF ≈ 2100 Hz



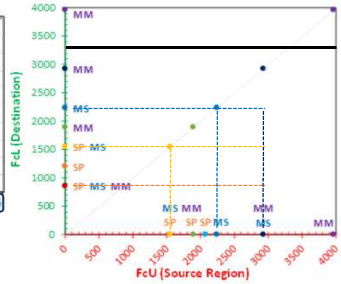
36

Results

VCVs



Moderately-Severe
MAOF ≈ 3300 Hz

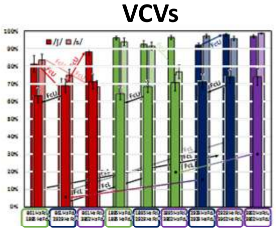


Depends on phoneme:
(↓ upper limit of source)
[ʃ], [s], /z/: ↑ % as FcU ↓

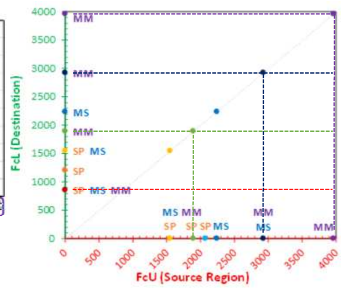
37

Results

VCVs



Mild-to-Moderate
MAOF ≈ 5000 Hz

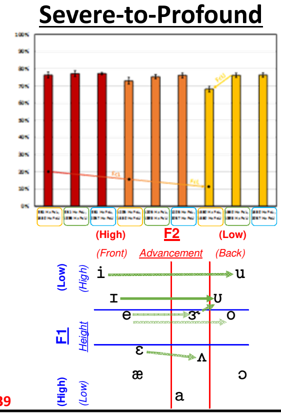


↑ % as FcL ↑
↑ % as FcU ↓
(higher FcL's, no effect)

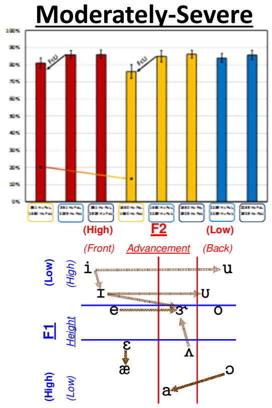
38

Vowels

Severe-to-Profound

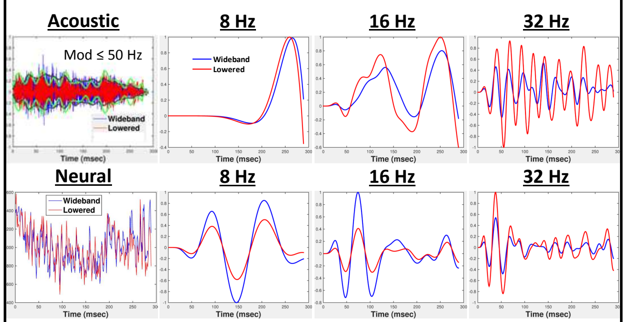


Moderately-Severe



39

Modulation Transfer Function Ratio (MTFR)



Zilany *et al*. (2014) model

40

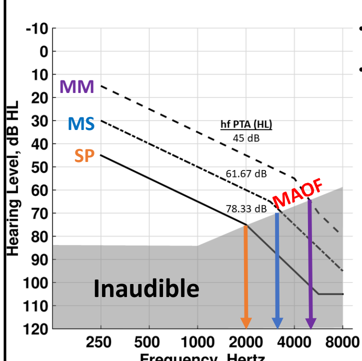
Potential Models/Mechanisms

	Stimuli	Max Input	Using /s/-/ʃ/ from s-sh Confusion Test					Using Test Stimuli	
			LF Edge	HF Edge	BW	A-MTFR	N-MTFR	A-MTFR	N-MTFR
SP	s-sh	-0.95	-0.95	-0.88	-0.52	0.97	0.97	—	—
	Fric	-0.88	-0.97	-0.73	-0.47	0.94	0.94	0.95	0.96
	VCV	-0.66	-0.85	-0.36	-0.29	0.73	0.74	0.84	0.82
MS	s-sh	-0.95	-0.89	-0.56	-0.89	0.96	0.96	—	—
	Fric	-0.82	-0.81	-0.75	-0.76	0.89	0.88	0.81	0.87
	VCV	-0.30	-0.19	-0.61	-0.11	0.58	0.38	0.81	0.92
MM	s-sh	-0.93	-0.62	0.21	-0.95	0.78	0.95	—	—
	Fric	-0.90	-0.70	0.55	-0.90	0.82	0.87	0.80	0.80
	VCV	-0.44	-0.49	0.78	-0.44	0.59	0.40	0.66	0.94

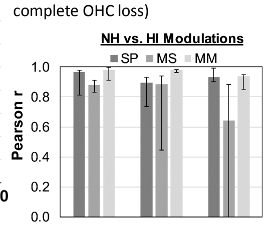
- Maximizing input freq. or spectral separation often led to significantly *lower* speech recognition
- *Narrower* output BW was sometimes associated with *higher* recognition→ ***Proxy for MTFR ???***
- High correlations with the MTFR indices implicates the **role of temporal modulation**

41

Auditory Nerve Model Simulations



- 8-channel WDRC hearing aid simulator (e.g., **Alexander, 2014**)
- **Zilany *et al*. (2014) model:** Low SR with OHC/IHC integrity parameters automatically selected by the AN model for each audiogram (SP had complete OHC loss)



42

Tested Assumptions

3. Maximizing [s]-[ʃ] edge frequency differences will maximize their discriminability
   - **FALSE,** this tends to decrease discriminability
     - Preserving envelope modulation maximizes discriminability
4. Maximizing [s]-[ʃ] discriminability will maximize perception of other fricatives and consonants
   - **TRUE** for fricatives, mostly true for consonants as a class
     - Settings that best preserve envelope modulations for these sounds also do the same for other high-frequency sounds

43

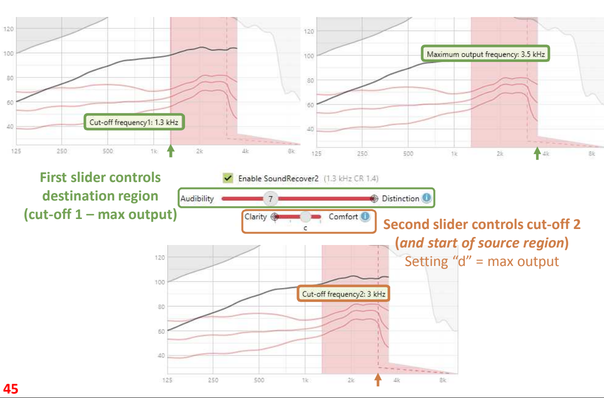
Future Directions

- ***Key point:*** *frequency-lowering methods and settings that preserve the greatest amount of temporal modulation from the original speech at the auditory periphery may yield the best outcomes for speech perception*

- Data collection
  - Collect lab data on the experiments described today (ANFC) using **hearing-impaired listeners**
    - Later, will collect field data using clinical devices
  - Pilot an **innovative variant of ANFC** that I have developed
  - Started collecting data on an innovative variant of **frequency transposition**

44

Programming Software Adjustments



45

SoundRecover2 Fitting Assistant

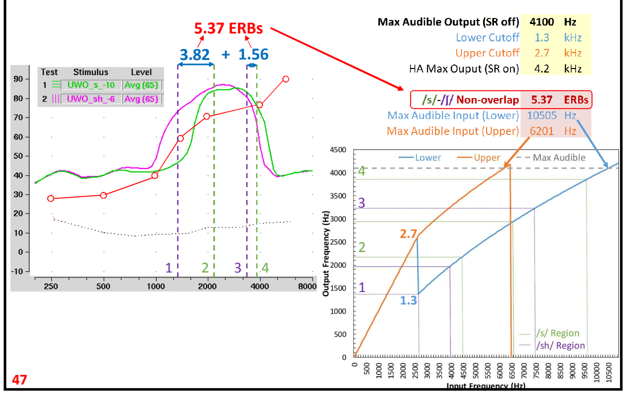
http://tinyURL.com/FLassist

- **Given that we know . . .**
  1. The frequencies of the edges of the calibrated UWO [s] and [ʃ] files
  2. The range of aided audibility (MAOF)
  3. The relationship between input and output frequencies for the different ANFC settings
     - *Cut-off 1, Cut-off 2, and Maximum Output frequencies*
- **Then, we can compute . . .**
  1. The amount of separation between the **high-frequency edges** of the UWO [s] and [ʃ]
  2. The amount of separation between the **low-frequency edges** of the UWO [s] and [ʃ]
  3. The **bandwidths** of the lowered UWO [s] and [ʃ]
  4. All of the above on a **psychophysical scale** that resembles normal **cochlear filtering (ERB scale)**

46

SoundRecover2 Fitting Assistant **v1.0**

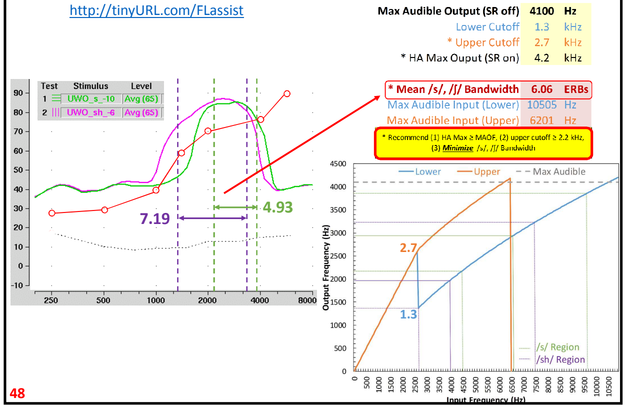
http://tinyURL.com/FLassist



47

SoundRecover2 Fitting Assistant **v2.0**

http://tinyURL.com/FLassist



48

Resources

Research Papers

- **Alexander, J. M.** (in revision). The S-SH confusion test and the effects of frequency lowering. Journal of Speech Language and Hearing Research.
- **Alexander, J.M.**, and Rallapalli, V., (2017). Acoustic and perceptual effects of amplitude and frequency compression on high-frequency speech. Journal of the Acoustical Society of America, 142, 908-923.
- Brennan, M.A., Lewis, D., McCreery, R., Kopun, J., and **Alexander, J.M**. (2017). Listening effort and speech recognition with frequency compression amplification for children and adults with hearing loss. Journal of the American Academy of Audiology, 28, 823-837.
- **Alexander, J.M.** (2016). Nonlinear frequency compression: Influence of start frequency and input bandwidth on consonant and vowel recognition. Journal of the Acoustical Society of America, 139, 938-957.
- Brennan, M.A., McCreery, R., Kopun, J., **Alexander, J.M.**, Lewis, D., and Stelmachowicz, P.G. (2016). Masking release in children with hearing loss when using amplification. Journal of Speech Language and Hearing Research, 59, 110-121.
- Rallapalli, V., and **Alexander, J.M.** (2015). Neural-Scaled Entropy predicts the effects of nonlinear frequency compression on speech perception. Journal of the Acoustical Society of America, 138, 3061-3072.
- **Alexander, J.M.**, Kopun, J.G., and Stelmachowicz, P.G. (2014). Effects of frequency compression and frequency transposition on fricative and affricate perception in listeners with normal hearing and mild to moderate hearing loss. Ear and Hearing, 35, 519-532.
- McCreery, R.W., **Alexander, J.M.**, Brennan, M.A., Hoover, B., Kopun, J., and Stelmachowicz, P.G. (2014). The influence of audibility on speech recognition with nonlinear frequency compression for children and adults with hearing loss. Ear and Hearing, 35, 440-447.
- Brennan, M.A., McCreery, R., Kopun, J., and **Alexander, J.M.**, Lewis, D., and Stelmachowicz, P.G. (2014). "Paired comparisons of nonlinear frequency compression, extended bandwidth, and restricted bandwidth hearing-aid processing for children and adults with hearing loss. Journal of the American Academy of Audiology, 25, 983-998.

50

Resources

Invited Review Papers

- **Alexander, J.M.** (in press). Frequency Compression and Transposition. In J. S. Damico & Martin J. Ball (eds.), The Sage Encyclopedia of Human Communication Sciences and Disorders.
- **Alexander, J.M.** (2016). 20Q: Frequency lowering ten years later – new technology innovations. *AudiologyOnline, Article #18040.*
- Angelo, K., **Alexander, J.M.**, Christiansen, T.U., and Jespersen, C.F. (2015). Oticon frequency composition. *Oticon White Paper*.
- **Alexander, J.M.** (2014). How to use probe microphone measures with frequency-lowering hearing aids. *Audiology Practices*, *6*, 8-13.
- **Alexander, J.M.** (2013). Individual variability in recognition of frequency-lowered speech. *Seminars in Hearing*, *34*, 86-109.
- **Alexander, J.M.** (2013). 20Q: The highs and lows of frequency lowering amplification. *AudiologyOnline*, *Article #11772*.
- Mueller, H.G., **Alexander, J.M.**, and Scollie, S. (2013). 20Q: Frequency lowering - the whole shebang. *AudiologyOnline*, *Article #11913*. *Rated one of the top AudiologyOnline articles of 2013*
- Mueller, H.G., **Alexander, J.M.**, and Scollie, S. (2013). Frequency lowering amplification: function, clinical applications, and practical tips. *AudiologyOnline*, *Article #23076*.

51