



Measuring spectral resolution for speech: Implications for cochlear implants

Matthew Winn

Ruth Litovsky

e-mail: mwinn2@wisc.edu

University of Wisconsin-Madison, USA



INTRODUCTION

Spectral resolution

Major limitation of listening with a cochlear implant

Important for distinguishing speech sounds

We want:

A way of measuring spectral resolution for speech sounds
* To see if it has been improved

Problem

Speech sounds vary by multiple acoustic dimensions (not just the spectrum)

Non-speech tasks (spectral ripples, pitch ranking) may bear limited relation to speech categorization

We are building an identification task

Everyday speech perception is identification, not discrimination

Our approach

Isolate and modify specific spectral cues found in natural speech [1] & measure listeners' identification of spectrally-driven contrasts.

Perception of spectral cues in speech sounds should:
... test for spectral perception **more directly** than word recognition tests
... **change** if we alter spectral resolution
... affect listeners' level of **effort** while listening

PARTICIPANTS

15 Listeners with bilateral cochlear implants

Code	Sex	Age	Years CI exp	Years BCI exp	Implant Type (Left)	Implant Type (Right)	Processor (Left)	Processor (Right)
IAJ*	F	67	16	9	Nucleus 24 Straight	Nucleus 24 Contour	N5	N5
IBF	F	62	7	5	Freedom Cont. Adv.	Freedom Cont. Adv.	Freedom	Freedom
IBK	M	72	9	3	Nucleus 24 Contour	Freedom Cont. Adv.	Freedom	N5
IBM	F	59	7	3	CI512 (N5)	Freedom Cont. Adv.	N5	N5
IBN*	M	66	13	3	Freedom Cont. Adv.	Nucleus 24 Contour	Freedom	Freedom
IBO	F	46	6	3	CI512 (N5)	Freedom Cont. Adv.	N5	Freedom
IBR	F	57	9	5	CI512 (N5)	Freedom Cont. Adv.	N5	Freedom
ICB	F	61	8	6	Freedom Cont. Adv.	Nucleus 24 Contour	Freedom	N5
ICD*	F	54	9	3	Freedom Cont. Adv.	Nuc. 24 Cont.	Freedom	Freedom
ICF	F	70	1	1	CI512 (N5)	CI512 (N5)	N5	N5
ICJ	F	63	1	3	CI512 (N5)	CI512 (N5)	N5	N5
ICM	F	59	1	1	CI512 (N5)	CI512 (N5)	N5	N5
ICO	F	32	1	1	CI512 (N5)	CI512 (N5)	N5	N5
ICP*	M	50	6	3	Freedom Cont. Adv.	Freedom Cont. Adv.	N5	N5
ICQ*	M	19	15	1	Freedom Cont. Adv.	Nucleus 24 Straight	N5	N5

* Indicates hearing loss occurred before age 5

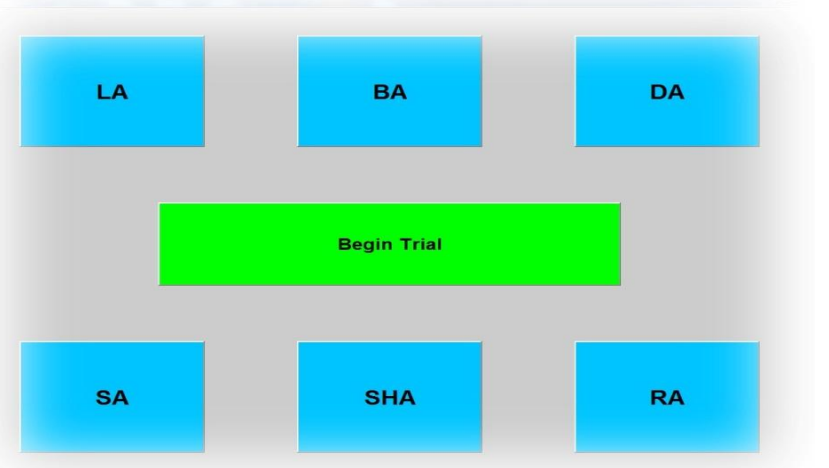
PROCEDURE

6-alternative forced choice

Target sounds: /ba/, /da/ ← contrasted by difficult spectral cue

Filler sounds:

/ja/, /sa/, /la/, /ra/ (for variability)



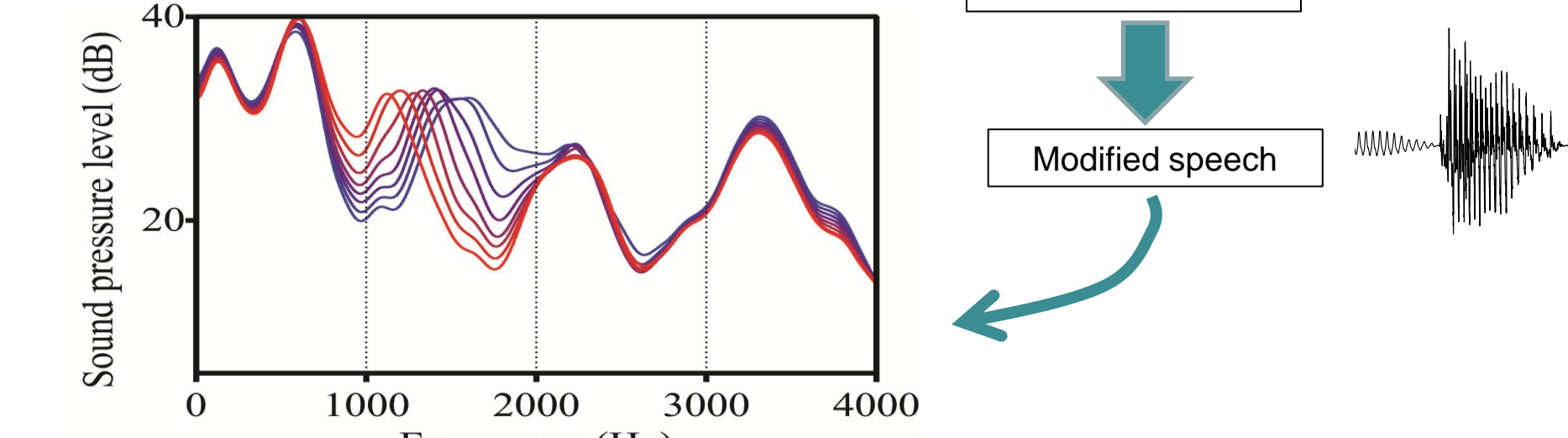
Interspersed with other listening tasks including open-set word recognition

HOW DO WE CREATE SPEECH STIMULI APPROPRIATE TO MEASURE SPECTRAL RESOLUTION?

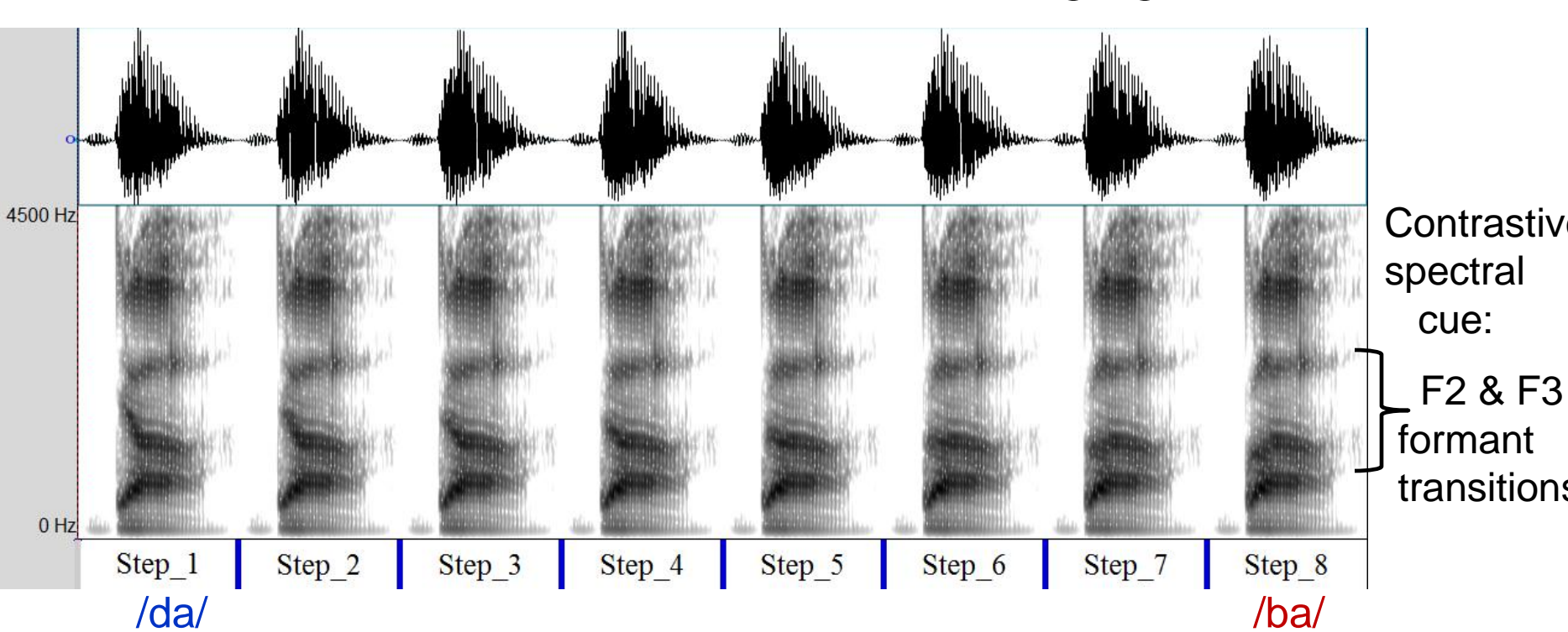
Natural utterances were decomposed into glottal source and vocal tract filter using linear predictive coding (LPC) inverse-filtering.

High-frequency energy (normally lost in LPC) was restored following re-synthesis.

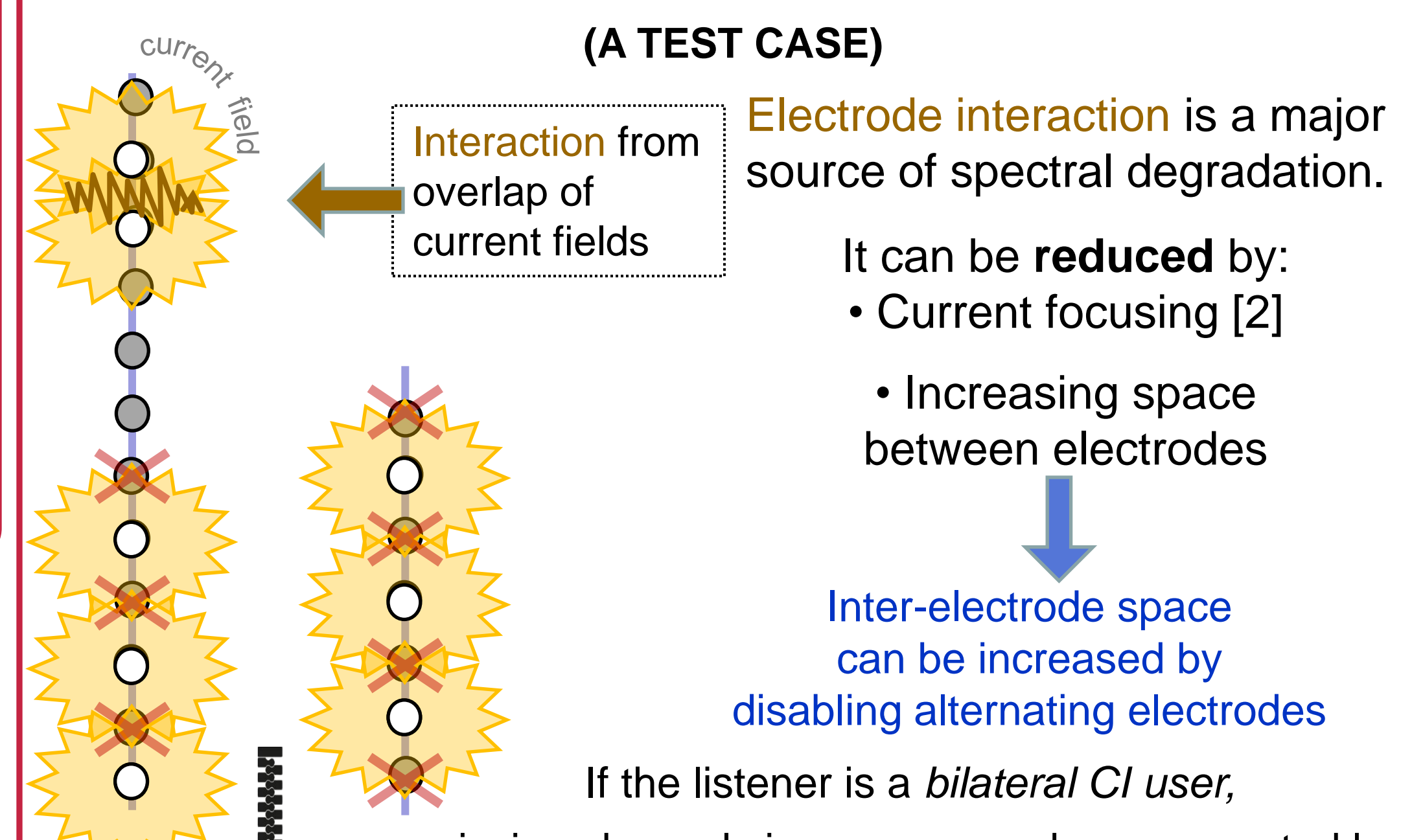
Stimuli maintained excellent natural quality with full bandwidth.



This process produced an 8-step continuum ranging from /da/ to /ba/.



HOW DID WE ATTEMPT TO IMPROVE SPECTRAL RESOLUTION? (A TEST CASE)



Electrode interaction is a major source of spectral degradation. It can be reduced by:
• Current focusing [2]
• Increasing space between electrodes
Inter-electrode space can be increased by disabling alternating electrodes

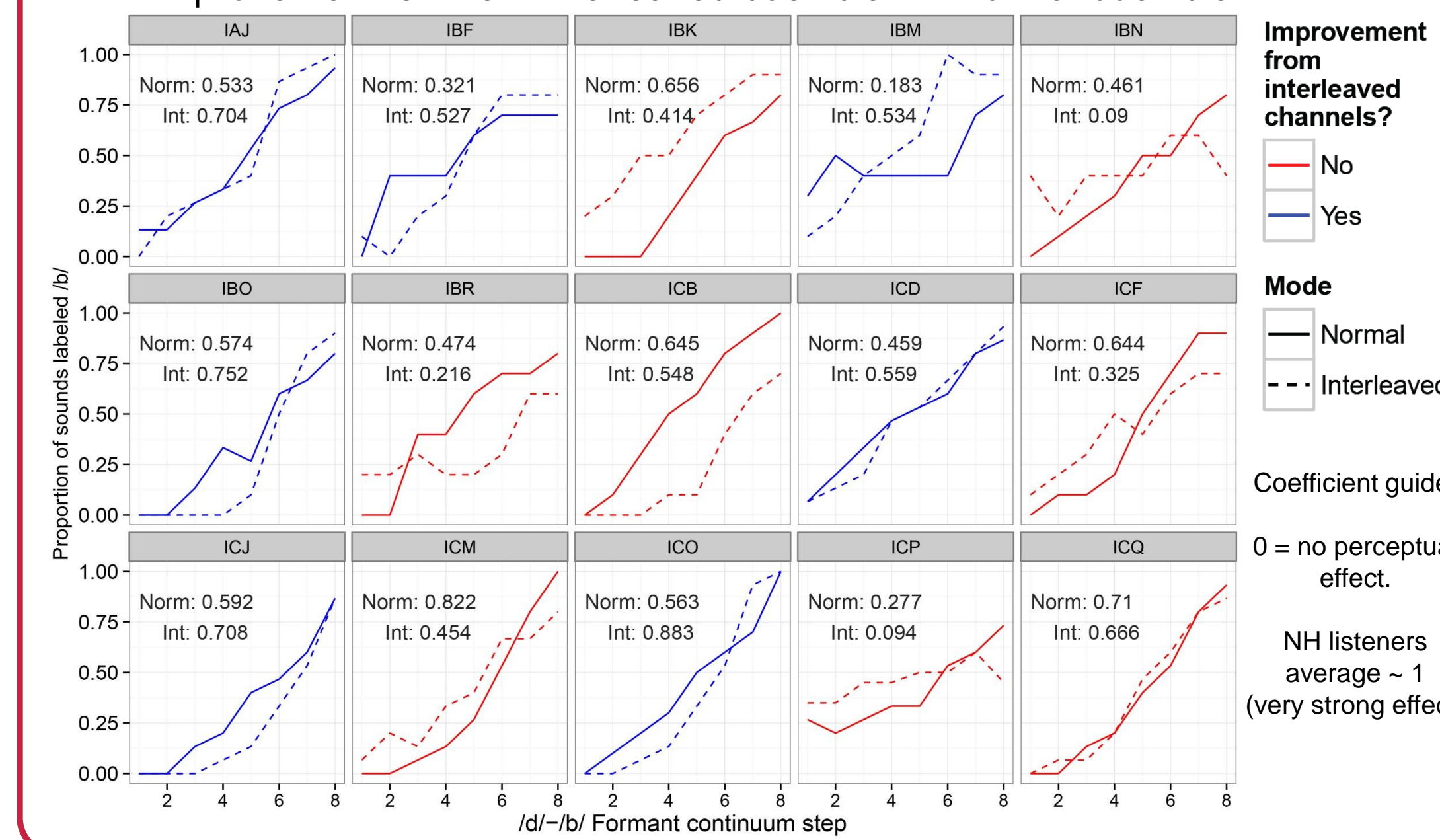
If the listener is a bilateral CI user, missing channels in one ear can be represented by active channels in the other ear*, creating deliberate non-overlapping spectral profiles in the two ears.

This is dichotic / interleaved / zippered channels [3,4,5,6]

* There is no guarantee that the interleaved channels will be aligned by pitch percept, but the listener suffers no loss of information input, compared to regular maps (i.e. no channels are lost)

ANALYSIS

1. /ba/-/da/ responses summarized for each continuum step to create sigmoidal psychometric functions
2. Functions modeled with logistic regression [7,8,9] using Laplace correction (maximum entropy prior).
3. The metric for spectral resolution is logistic coefficient for the spectral cue (similar to "cue-weighting" analyses)
4. Coefficients for normal & interleaved conditions are compared. Improvement is when: interleaved coefficient > normal coefficient



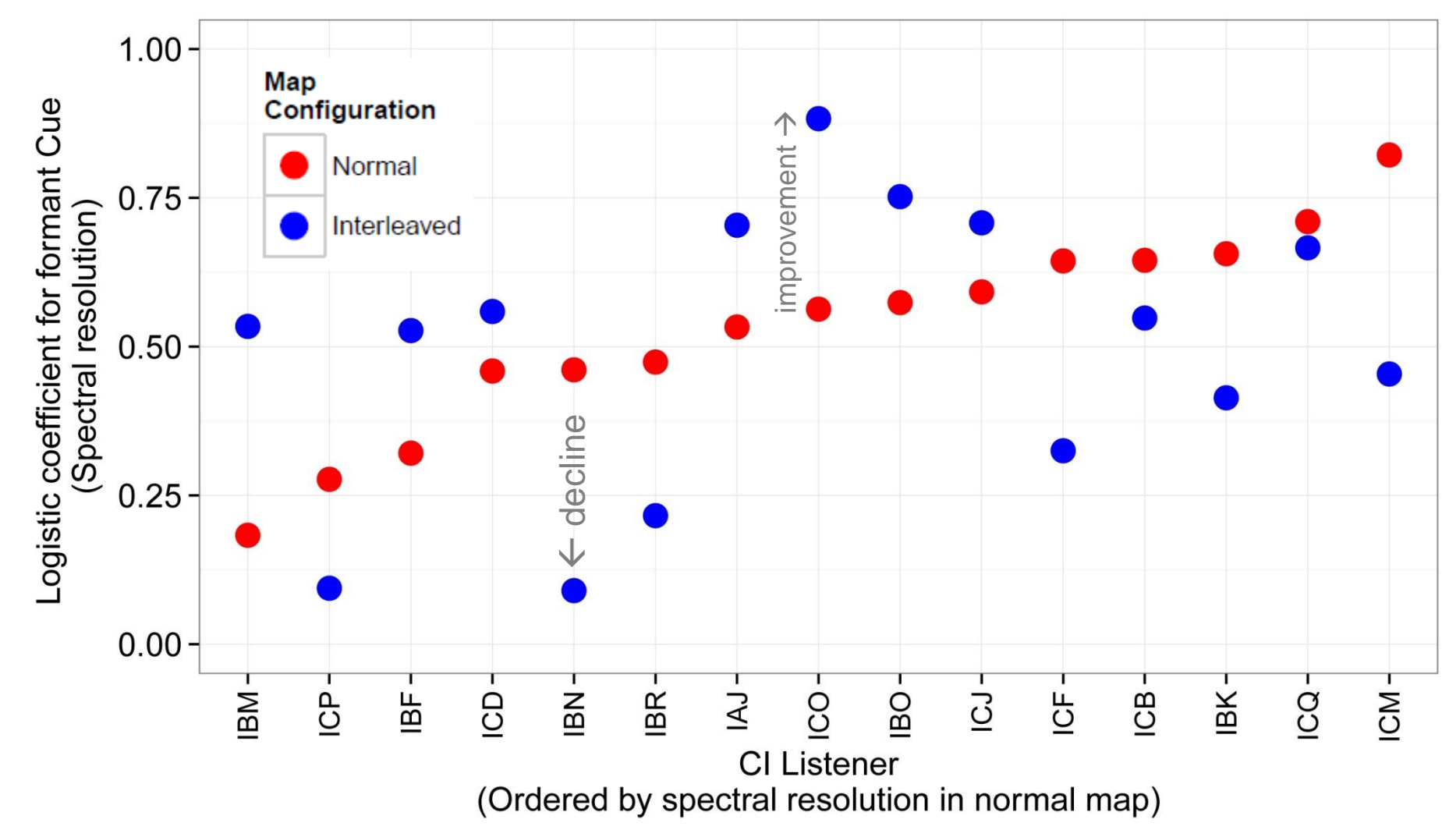
RESULTS

Spectral Resolution

Formant coefficient = log odds of perceptual change resulting from a change of the formant cue (higher number → better use of the spectral cue)
NH listeners → average around 1.0

Listeners showed variability in resolution with their basic maps (wide range of red values)

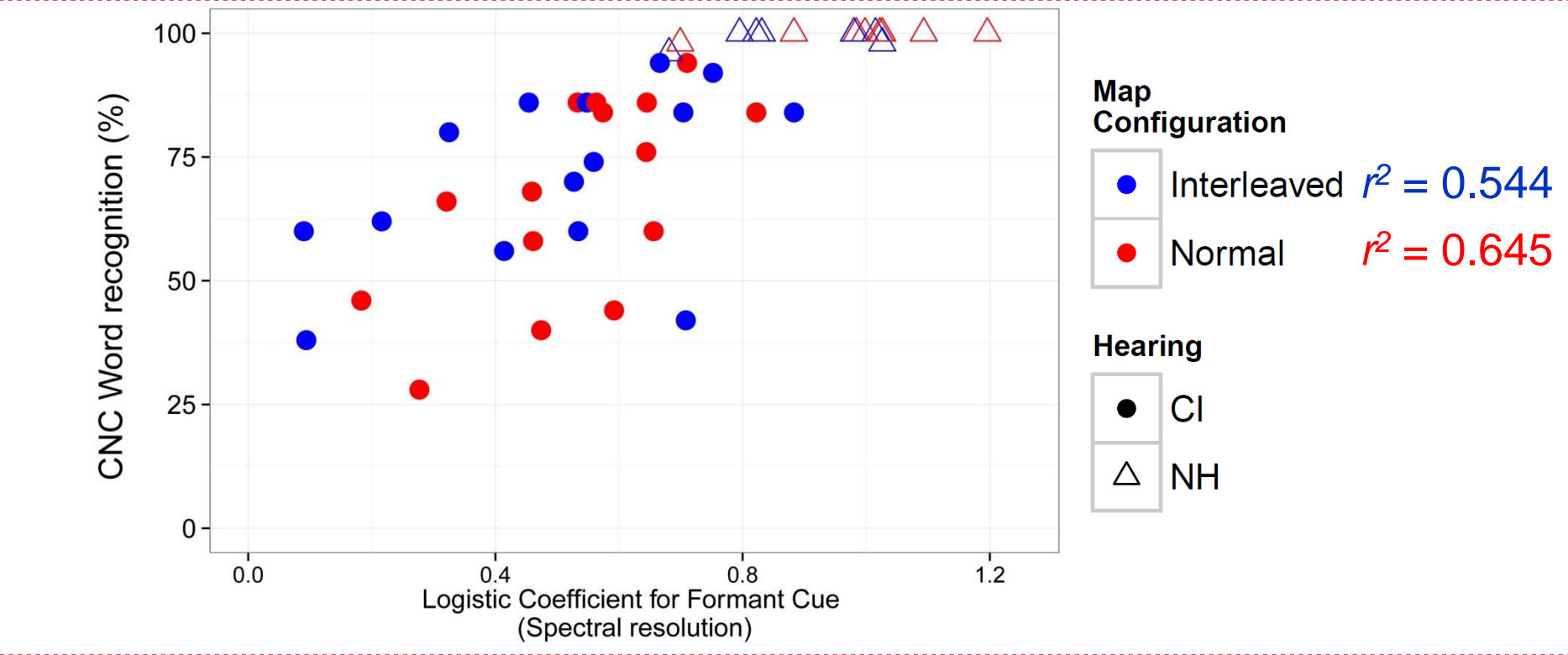
7 listeners showed improvement with interleaved maps (blue above red), while 8 showed decline (blue below red)



Spectral resolution and word recognition

Listeners with better spectral resolution tend to score higher on word recognition.

However, variance in word recognition is also explained by other factors (temporal resolution, etc.)

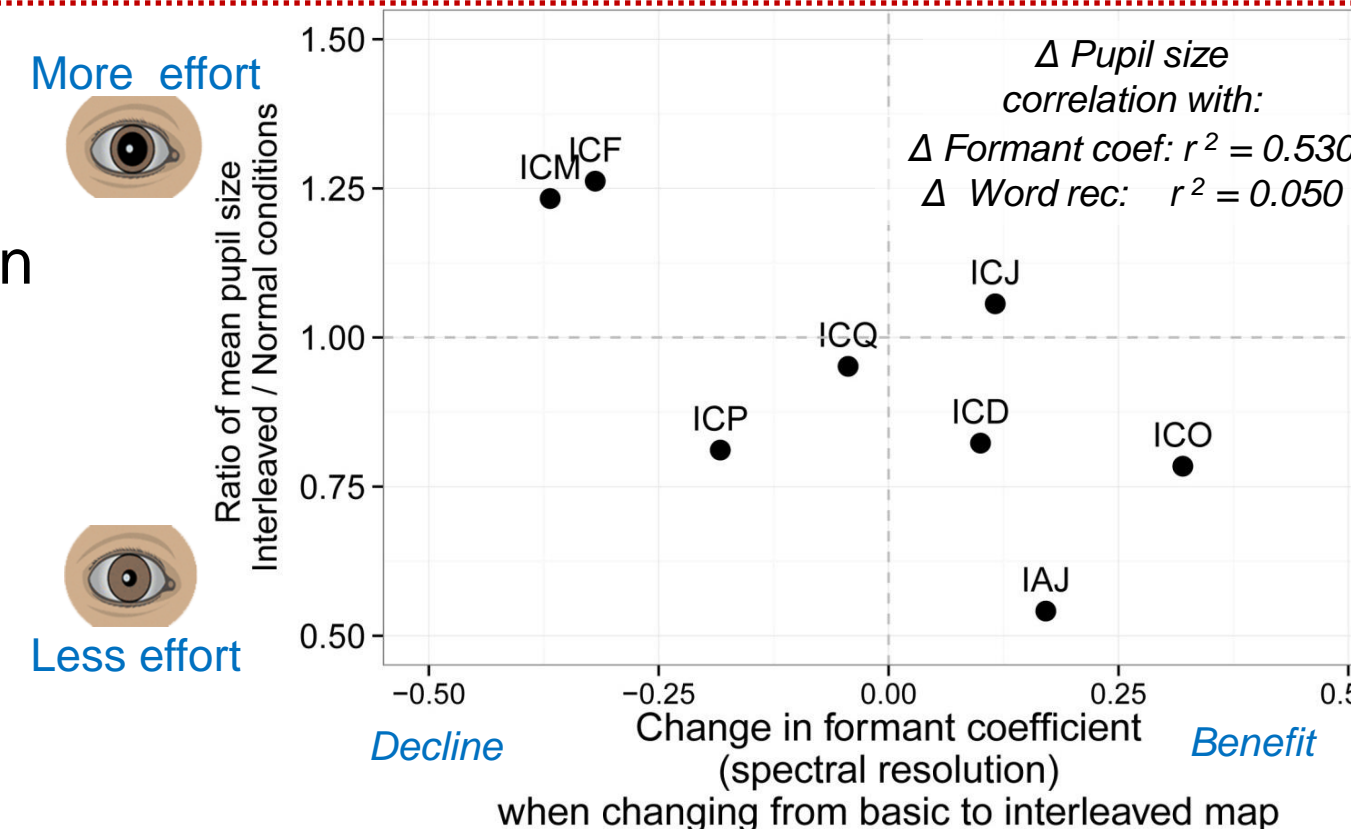


Spectral resolution and listening effort

Quantified using pupil dilation during an IEEE sentence repetition task [10]
Pupil size measured with a Tobii 2150 eye-tracker between sentence & repetition

Listeners who benefited from interleaved channels also showed benefit in terms of listening effort as indicated by less pupil dilation while listening with interleaved channels.

Change in effort with interleaved channels was better predicted by the spectral resolution test than by word recognition.



CONCLUSIONS

- ★ Spectral resolution can be measured using speech stimuli that are explicitly controlled in the spectral domain - Accurate and robust use of spectral cues in speech is the operational metric for spectral resolution
- ★ Performance with these stimuli is related to open-set word recognition
- ★ This method of testing could be used to evaluate ways to improve spectral resolution (e.g. current focusing)
- ★ Improvement in spectral resolution can be validated using tests of listening effort, including pupillometry - Improving spectral resolution generally results in less pupil dilation, which is a sign of reduced listening effort.

[1] Boersma, P. & Weenink, D. (2013). Praat: doing phonetics by computer [Computer program]. Version 5.3.41
 [2] Landsberger, D. et al. (2012). Reducing current spread using current focusing in cochlear implant users. Hearing Research, 284, p. 16-24.
 [3] Mani, A. et al. (2004). Dichotic speech recognition by bilateral cochlear implant users. Int'l Congress Series, 1273, p. 466-469.
 [4] Loizou P. et al. (2003). Dichotic speech recognition in noise using reduced spectral cues. JASA, 114, 475-483.
 [5] Tyler, R. et al., (2010). An attempt to improve bilateral cochlear implants by increasing the distance between electrodes and providing complementary information to the two ears. JAAA, 21, p. 52-65.
 [6] Zhou, N. (2010). Lexical Tone Development, Music Perception and Speech Perception in Noise with Cochlear Implants: The Effects of Spectral Resolution and Spectral Mismatch. Doctoral dissertation
 [7] R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.
 [8] Bates, D. & Maechler (2013) lme4: Linear mixed-effects models using Eigen and R syntax. R package Version 0.999999-2
 [9] Morrison, G. & Kondaurova, M. (2009). Analysis of categorical response data: Use logistic regression rather than endpoint-difference scores or discriminant analysis. JASA, 126, p. 2159-2162.
 [10] Koelewijn, T. et al. (2012). Pupil dilation uncovers extra listening effort in the presence of a single-talker masker. Ear & Hearing, 33, p. 291-300.