HEARING_Pontoppidan_Vatti.ai - 10.01.2014 - 23:02:24

Temporal fine structure deficits in perceptual organization

Niels Henrik Pontoppidan, Oticon Research Centre Eriksholm & Marianna Vatti, Centre for Applied Hearing Research, Technical University of Denmark.

Introduction

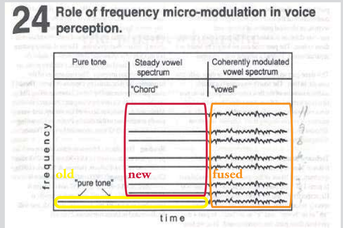
Recent experiments show that hearing-impaired listeners have limited access to temporal fine structure (TFS) [Hopkins & Moore, 2008; Lunner 2009] while others indicate that normal-hearing listeners benefit from TFS in difficult situations [Lunner et. al. 2010; Hopkins, Moore, and Stone 2009, Lorenzi 2006].

In the TFS1 test [Moore and Sek 2009], replicated at Eriksholm most hearing impaired cannot perceive the seemingly large differences between the harmonic and the shifted tones; this is despite the fact that all involved parts are audible and that differences are obvious to normal-hearing listeners.

The observations raise the question: What else can hearing impaired not perceive when they cannot discriminate between the TFS1 test stimuli?

Micro-modulation

Chowning's *Role of Micro-modulations* from 1979 [Bregman 1990, Bregman and Ahad 1996] demonstrates how the onset of relatively small amount of frequency modulation (the micro-modulation) alters the perceptual organization of the sound.



The effect extends to vowel mixtures, where independent micro-modulation in three different vowels at 400, 500, and 600 Hz (g,b,d) segregate after the micro-modulation onset.

Hearing impairment

Impaired micro-modulation perception could have devastating consequences for both auditory scene analysis and understanding speech. However, the extent of micro-modulation perception impairment must be investigated. Moreover, the consequences of that impairment must be investigated, especially since people suffering from Amusia do not seem to have problems with understanding speech [Peretz, 2009]. If it turns out that hearing-impaired cannot perceive and group according to the micro-modulation, it strengthens the claim that TFS is important for speaker recognition and gap listening.

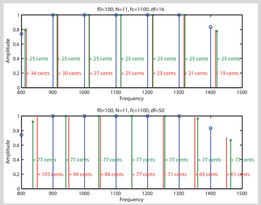
TFS1 test stimuli and pitch

The shifted TFS1 stimuli are created by moving all partials of a harmonic complexes df Hz upwards.

The authors perceive the shifted TFS1 stimuli as having higher pitch than the TFS1 reference - it is the cue that we remember and refer to. The lack of harmonic structure only seems to play a perceptual role for very low harmonics. Furthermore, shifted TFS1 stimuli resemble harmonic complexes with detuned partials for reasonable choices of N and df.

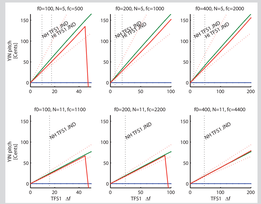
$$\text{Cents}=1200\ \log_2(1+df/f_{ref})$$

The line spectra to the right displays the large similarity between the shifted TFS1 test stimulus (red) and a (normal) harmonic complexes with pitch f0+df/N (green). The TFS1 references is blue. The shift in cents for the TFS1 and harmonic with respect to the corresponding partial in the reference stimuli is denoted as well.



Pitch estimation

Analysing the noise free TFS1 stimuli with YIN reveals a similar pattern. For a resolved central component (N=5) YIN consistently estimates the pitch as being lower than f0+df/N. For N=11, YIN's pitch estimate follow f0+df/N.



Moreover, the raised pitch f0+df/N is consistent with Santurette's pitch matching results with similar stimuli [Santurette and Dau, 2010].

Micro-modulation and TFS1 test

In the case of resolved partials (N=5) the hearing-impaired test persons can identify shifts down to 30-40 cents. However, for N=11, most of the hearing-impaired test persons could not identify the interval with the maximal TFS1 df shift corresponding to 77 cents (or 105 for the lowest partial).

Since YIN estimates the amount of micro-modulation in Chowning's vowels to be ± 50 cents, it could indicate that hearing-impaired listeners wouldn't perceive the shared micro-modulation in all partials of Chowning's stimuli, and therefore wouldn't group accordingly.

However, we have not found the right questions that could accompany the stimuli in a detection test.

Micro-modulation and vowel perception

Synthetic vowels sound mechanical and unnatural if micro-modulations are not present. At the same time naturalness decreases if the frequency modulation gets too fast or too deep. Investigations by Sundberg, McAdams, de Cheveigné as well as many others have indicated that the micro-modulation *sweet spot* of is in the range 3 to 8 Hz. Moreover, the micro-modulation also include a noisy component with low bandwidth.

Does the sweet spot vary with hearing impairment?

Experiment designed to infer peoples preferred modulation rate and depth.

Stimuli: 400 Hz Harmonic complex with spectrum shape like Chowning's 400 Hz vowel.

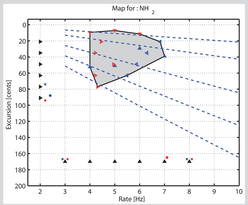
Procedure: Listeners are asked to judge whether the vibrato in a synthetic vowel sounds natural answering yes or no.

Each condition require two measurements, one starting from the low extreme, and one starting from the high extreme.

The conditions either measure the preferred :
- modulation depth for a fixed modulation rate (↳↑), or
- modulation rate for a fixed modulation depth (→↲).

A 1u-1d adaptation-rule is used to converge to the threshold.

The figure shows one listener's results from testing 6 depths and 6 rates from both extremes. The blue and red triangles show the found boundaries.

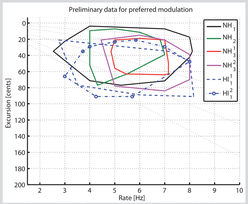


The listener did not rate any of the examples with 3 Hz, 8, and 90 cents as being natural.

The preferred modulation range is a convex hull including the modulation rate and depth boundaries. The dashed blue lines indicate iso-modulation-index curves, that might influence the lower depth boundary.

Preliminary results

The figure shows the preferred modulation range for a few test persons (including the authors). Some test-retest variability is indicated by the two listeners with repeated measures.



The lower boundary for the single hearing-impaired listener seems to be close to 20 cents, the minimal FMDL for hearing-impaired reported by Moore and Skrodzka. This follows the expectation that the listener judged the naturalness of the modulation using the low harmonics.

Two of the normal hearing listeners prefer modulation considerably below Moore and Skrodzka's thresholds for hearing-impaired.

Vowel identification

The link between micro-modulation and vowel identification has been investigated numerous times. Marin and McAdams [1991] indicated that micro-modulation was more important for the vowel quality than for vowel identification. Moreover, other attempts have shown that the micro-modulation might decrease vowel identification scores [Sundberg, 1975].

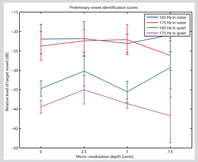
Nevertheless, acknowledging that the importance of TFS seems to be in difficult situations [Lunner et al., 2010] and vanishes in easy, it could be argued that previous attempts at demonstrating the link between micro-modulation and vowel identification were based on to easy tasks - or had too many other cues.

Stimuli: Three "Danish" vowels 'a', 'i', and 'o' synthesized using Klattgrid [Wienink, 2010]. The vowels specified by Marin and McAdms [1991] where made Danish by substituting the first two formants with the formant frequencies sugguested by Grønnum [2005]. Vowels with 156 Hz pitch where used as maskers, while vowels with 165 and 175 Hz pitch where used as targets.

Conditions: Two simultaneous vowels with and without TEN noise at the same level as the masker vowel. The micro-modulation depth in the target vowel was varied.

Procedure: Listener is asked to identify the vowel with the higher pitch. The 2D-1U adaptive tracks the level of the target vowel.

Preliminary results: vowel identification performed by author NHP have not indicated any benefit of micro-modulation in the identification of vowels. The seemingly trend from 5 cents to 7,5 cents was not observed in similar runs with 10 and 20 cents depth.



Potential problems with stimuli:

1) 'a', 'i', and 'o' are the end points of the vowel triangle [Grønnum, 2005] and thus relatively easy to keep apart. This could render TFS cues useless in the specific task.

2) The 'i' and 'o' could be more natural. Especially important in the quiet condition where 'i' can be identified by some high frequency rating.

Wrap up

The perceptual organization may be influenced by the lack of TFS sensitivity.

The minimal preferred micro-modulation depth may be larger for hearing-impaired than for normal hearing - perhaps due to increased detection thresholds.

TFS1 thresholds and FMDL's indicate that lack of TFS sensitivity deprive hearing-impaired of some of the small frequency modulation depths that normal-hearing do find natural.

No indication of benefits from micro-modulation in identification of synthetic vowels in quiet and in TEN noise.

Contact information

Niels Henrik Pontoppidan - nhp@oticon.dk
Oticon Research Centre Eriksholm
Kongevejen 243, DK-3070 Snekkersten, Denmark

References

Bregman, A.S. *Auditory Scene Analysis*. MIT Press, Cambridge, Massachusetts 1990.
Bregman, A.S. and Ahad, P.A. *Demonstrations of Auditory Scene Analysis: The Perceptual Organization of Sound*. Compact Disc, 1996. MIT Press.
de Cheveigné, A. and Kawahara, H. *YIN, a fundamental frequency estimator for speech and music*. JASA 111:1917, 2002.
Grønnum, N. *Fonetik & Fonologi*, 3rd Edition, Akademisk Forlag, 2005
Hopkins, K. and Moore, B.C.J. *Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information*. JASA 122(2):1055-1068. 2007.
Hopkins, K., Moore, B.C.J., and Stone, M.A. *Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech*. JASA 123(2):1140-1153, 2008.
Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., Moore, B.C.J. *Speech perception problems of the hearing impaired reflect inability to use temporal fine structure*. PNAS 103(39):18866-18869, 2006.
Lunner, T., Hietkamp, R.K., Andersen, M.R., Hopkins, K., and Moore, B.C.J. *Effect of speech material on the benefit of temporal fine structure information in speech for normal-hearing and hearing-impaired subjects*. Submitted to JASA 2010.
Marin, C.M.H. and McAdams, S. *Segregation of concurrent sounds. II: Effects of spectral envelope tracing, frequency modulation coherence, and frequency modulation width*. JASA 89:341, 1991.
McAdams, S.E. *Spectral fusion, spectral parsing and the formation of auditory images*, PhD Thesis, Stanford University, 1984
Moore, B.C.J. and Skrodzka, E. *Detection of frequency modulation by hearing-impaired listeners: Effects of carrier frequency, modulation rate, and added amplitude modulation*. JASA 111(1):327-335, 2002.
Moore, B. and Sek, A. *Development of a fast method for determining sensitivity to temporal fine structure*. International Journal of Audiology 48(4):161-171, 2009.
Peretz, I. *Abnormal connectivity in the auditory-frontal neural pathway in congenital pitch disorder*. Workshop "Auditory temporal processing in normal and impaired ears". 2009.
Santurette, S. and Dau, T. *The low pitch of high-frequency complex tones relies on temporal fine structure information*. JASA (submitted) 2010.
Sundberg, J. *Vibrato and vowel identification*. STL-QPSR 16(2-3), 1975

Acknowledgements

The design of the vowel perception study was greatly improved by inputs from Sébastien Santurette, Torsten Dau, Renskje Hietkamp, and Thomas Lunner.

CAHR Centre for Applied Hearing Research

oticon Research Centre Eriksholm