REFERENCES

Arehart, K. H., Kates, J. M., Anderson, M. C., and Harvey, L. O. (**2007**). "Effects of noise and distortion on speech quality judgments in normal-hearing and hearing-impaired listeners" J. Acoust. Soc. Am., **122** (2), 1150-1164.

Bentler, R., Wu, Y.-H., Kettel, J., and Hurtig, R. (**2008**). "Digital noise reduction: Outcomes from laboratory and field studies" Int. J. Audiol., **47** (8), 447-460.

Bjerg, A. P., and Larsen, J. N. (**2006**). *Recording of Natural Sounds for Hearing Aid Measurements and Fitting*, Ørsted, Denmark: Danish Technical University (DTU), Acoustic Technology.

Brons, I., Houben, R., and Dreschler, W. A. (**2011**). "Perceptual effects of noise reduction in hearing aids" in ISAAR - International Symposium on Auditory and Audiological Research.

Dreschler, W. A., Verschuure, H., Ludvigsen, C., and Westermann, S. (**2001**). "ICRA noises: artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment" International Collegium for Rehabilitative Audiology. *Audiology: official organ of the International Society of Audiology*, **40** (3), p.148.

Hoetink, A. E., Körössy, L., and Dreschler, W. A. (**2009**). "Classification of steady state gain reduction produced by amplitude modulation based noise reduction in digital hearing aids" International Journal of Audiology, **48** (7), 444–455.

Houben, R., Dijkstra, T. M. H., and Dreschler, W. A. (**2011**). "Differences in preference for noise reduction strength between individual listeners" in *130th Convention of the Audio Engineering Society*. London, pp. 1-9.

Loizou, P.C., and Kim, G. (**2011**). "Reasons why current speech-enhancement algorithms do not improve speech intelligibility and suggested solutions" IEEE Transactions on Audio, Speech, and Language Processing, **19** (1), 47–56.

Luts, H., Eneman, K., Wouters, J., Schulte, M., Vormann, M., *et al.* (**2010**). "Multicenter evaluation of signal enhancement algorithms for hearing aids" J. Acoust. Soc. Am., **127** (3), 1491-1505.

Versfeld, N. J., Daalder, L., Festen, J. M., and Houtgast, T. (**2000**). "Method for the selection of sentence materials for efficient measurement of the speech reception threshold" J. Acoust. Soc. Am., **107**(3), 1671-1684.

Wickelmaier, F., and Choisel, S., (**2006**). "Modeling within-pair order effects in paired-comparison judgments" in *22nd Annual Meeting of the International Society for Psychophysics*. St. Albans, Hertfordshire, England, pp. 89–94.

Zhao, D. Y., Kleijn, W. B., Ypma, A., and de Vries, B., (**2008**). "Online Noise Estimation Using Stochastic-Gain HMM for Speech Enhancement" IEEE Transactions on Audio Speech and Language Processing, **16**(4), p.835.

Fast and intuitive methods for characterizing hearing loss

DIRK OETTING[1,2], BIRGER KOLLMEIER[2] AND STEPHAN D. EWERT[2]

[1] *Fraunhofer Institute for Digital Media Technology, Oldenburg, Germany*

[2] *Medizinische Physik, Carl-von-Ossietzky University Oldenburg, Germany*

The possibility of integrating hearing-aid technology like dynamic compression in current and future consumer audio devices raises the question how parameters of hearing supportive algorithms can be adjusted by the user to either compensate the individual hearing loss or to accommodate listening preferences. Here, three methods for measuring the auditory capacity based on loudness judgments and comparisons were evaluated. All methods used a simple interface and appear generally suited for integration in consumer audio electronics. Results of the suggested methods were compared to adaptive categorical loudness scaling [ACALOS, Brand and Hohmann, J. Acoust. Soc. Am. 112, 1597-1604 (2002)]. Gain prescriptions were derived for narrow-band loudness compensation based on the suggested methods, the clinically applicable ACALOS, and for NAL-NL2 [Keidser and Dillon, Hearing Care for Adults, 133-142 (2006)]. All loudness based procedures led to similar gains.

INTRODUCTION

Less than 20% of the mild-to-moderate hearing impaired population uses a hearing aid (Hougaard and Ruf, 2011) although the majority would benefit from hearing supportive technologies. To overcome stigma particularly for groups with mild hearing loss, hearing supportive technology could be integrated into communication and media devices (e.g., mobile phones, TVs, music players). These devices offer sufficient signal-processing power and capability to deliver high-fidelity sound quality, however, the problem of the individual fitting is still unsolved.

Standard audiometric measurements like hearing threshold appear not suited for integration in un-calibrated audio products, particularly if used in noisy environments. Moreover, like for hearing-aid fitting, knowledge about supra-threshold hearing deficits might be beneficial. Here, three fast and intuitively accessible methods, motivated by the adjustment un-calibrated video monitors based on video test images are suggested.

Instead of adjusting screen luminance levels to achieve well separable brightness impressions, sound levels were adjusted to match the well separable loudness categories “just audible”, “soft”, “comfortable”, and “loud”. The adjustment was either independent in three different frequency regions or additional loudness comparisons across frequency were included. The results were compared to laboratory measurements of the audiogram and adaptive categorical loudness scaling (ACALOS) for normal-hearing and hearing-impaired listeners.

Proceedings of ISAAR 2011: Speech perception and auditory disorders. 3rd International Symposium on Auditory and Audiological Research. August 2011, Nyborg, Denmark. Edited by T. Dau, M. L. Jepsen, J. Cristensen-Dalsgaard, and T. Poulsen. ISBN 87-990013-3-0. EAN 9788799001330. The Danavox Jubilee Foundation, 2012.

METHODS

Subjects

10 normal-hearing (NH) and 11 hearing-impaired (HI) listeners participated in this study. The NH group was aged between 21 and 65 years (mean: 32.9 years, standard deviation: 12.0 years) and had audiometric thresholds of 20 dB HL or better at the test frequencies of 500, 2000, and 6000 Hz (ANSI, 2004). Their mean results were used as reference values for the HI group. The HI group was aged between 25 and 73 years (mean: 60.5 years, standard deviation: 18.0 years) with mild-to-moderate sensorineural hearing losses (PTA of 18.3 – 55.0 dB). A large range of typical hearing losses was covered: 7 cases of sloping hearing loss, 2 times steep sloping hearing loss, and 2 cases of u-shaped hearing loss. Subjects received an hourly compensation for taking part in the study.

Procedures

Three different loudness-based "auditory test image" methods for measuring the auditory capacity were developed fitting the following requirements for consumer products: (a) uncomfortably loud sounds are avoided, (b) only a few feedback buttons are available, (c) no graphical display is required. Narrow-band noise (1/3 octave) centered at 500, 2000, and 6000 Hz was used in all methods. Subjects had to adjust the volume of the stimuli to the loudness categories "soft", "comfortable", and "loud". The user interface consisted of three feedback buttons ("plus", "minus", and "next"). The plus button increased the signal level by 3 dB whereas the minus button reduced it by the same amount. After each level modification, the stimulus was presented to the listener. If the listener was satisfied with the setting, the next button led to the next condition.

Method 1: Across frequency

Starting at 2000 Hz, the listeners had to adjust the level to a comfortable loudness perception. In a matching task, the levels of the 500 and 6000 Hz noises were than adjusted to the same loudness as perceived for the 2000-Hz noise which was played as a reference at the previously adjusted comfortable level. The same procedure was then repeated for the loudness categories soft and loud.

Method 2: Across levels

In this method, the auditory capacity was measured independently at each of the three frequencies. The level of the narrow-band noise had to be adjusted to comfortable, loud and soft. After adjustment for one (center) frequency, a control stimulus containing all levels in increasing order was played and the listener was asked whether the three levels matched the desired loudness categories. If not, re-adjustment with the same procedure was possible, starting from the already adjusted levels. This procedure was repeated for all frequencies and did not contain loudness matching across frequencies.

Method 3: Combined method

In method 3, the two previous methods were combined. Starting at 2000 Hz, the levels had to be adjusted to comfortable, soft and loud. The same control stimulus as described in method 2 was then presented and re-adjustment was allowed. After all three levels were set at 2000 Hz, the listener had to adjust the levels for 500 and 6000 Hz in the loudness matching task from method 1 in the order comfortable, soft and loud.

Finally, in each of the methods 1-3, the narrow-band noises had to be set to a just audible level allowing for a comparison with threshold measurements.

Adaptive categorical loudness scaling (ACALOS)

For comparison, loudness scaling using the same narrow-band noise stimuli was performed. The stimuli were presented at different sound levels in randomized order and the listeners had to rate the perceived loudness on a categorical scale from "inaudible" to "too loud" (for details see Brand and Hohmann, 2002). Figure 1 shows the typical shape of the resulting loudness perception functions of a NH and HI listener.

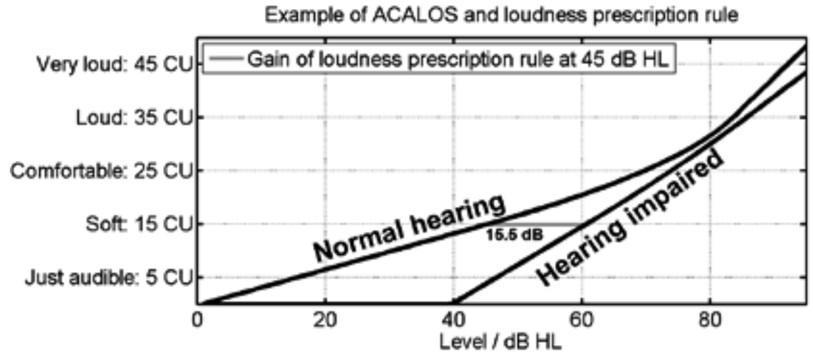

Fig. 1: Example of ACALOS loudness scaling for a normal-hearing and hearing-impaired listener. The HI listener shows a steepened loudness function known as loudness recruitment. The red line indicates the gain necessary to restore the normal-hearing loudness function for narrow-band noises (narrow-band loudness compensation).

Adaptive threshold measurements

Two pure tones were used to measure the hearing threshold according to the single interval method suggested by Lecluyse and Meddis (2009). The level of the first (cue) tone was 10 dB higher than the level of the target tone. The subject was asked to indicate the number of certainly perceived tones. A 1-up, 1-down alternative-forced choice (AFC) procedure was used to reduce or increase the level of both tones if two or one tone was heard, respectively. Step sizes were 10 dB initially and 2 dB after the first reversal. The procedure finished 10 trials after the first reversal.

The initial catch trial rate was 20%. The rate was increased by 10% on the next run (max. value 50%) if the listener was caught giving a false detection..

Audiometric measurements

Air-conduction measurements were made with the Siemens UNITY 2 audiometer connected with Sennheiser HDA200 circumaural headphones. Listeners had to raise their hand to indicate whether they perceived the tone. Continuous pure-tones signals at 125, 250, 500, 750, 1000, 1500, 2000, 3000, 4000, 6000, and 8000 Hz were used. In the repetition measurements, the audiometric measurements were only carried out at 500, 1000, 2000, and 6000 Hz. Starting at -10 dB HL, the level was increased by 5 dB until the first response occurred. The level was decreased by 15 dB to repeat the measurement. The threshold was defined as the lowest level where two responses out of three occurred.

Lab measurements and schedule

Four repetitions of each of the above described measurements were conducted in three sessions. Data for each method were collected on two different days.

1. Audiogram, single-interval threshold, method (1,2,3), method (2,3,1), audiogram
2. Audiogram, method (3,1,2), method (3,2,1), 2xACALOS, single-interval threshold
3. Audiogram, 2xACALOS, single-interval threshold

RESULTS

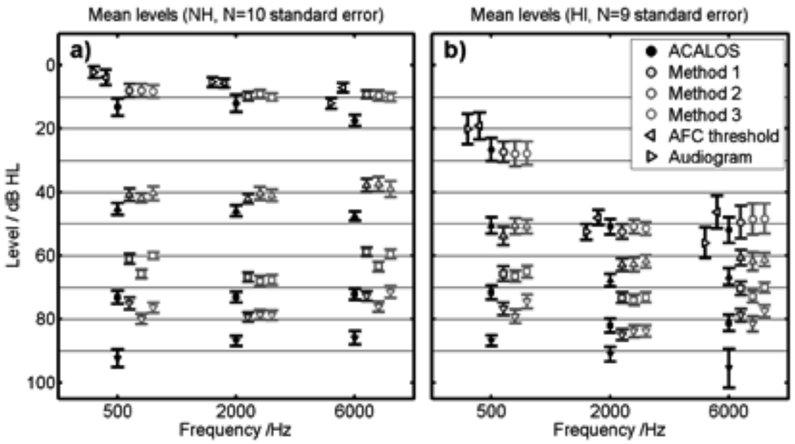

Fig. 2: Mean levels and standard errors for normal hearing a) and hearing impaired listeners b) for audiogram and AFC (single-interval) threshold. Additionally, the levels for just audible, soft, comfortable, and loud for method 1-3 and from ACALOS are shown.

In Fig. 2, the mean values and the standard error for normal-hearing and hearing-impaired listeners are shown. The results of two hearing-impaired subjects have been excluded because they reached the upper limit of 95 dB HL for adjusting the category loud at 6000 Hz. Levels for soft were about 40 dB HL, comfortable levels were in the range of 60-75 dB HL, and loud levels were found to be above 75 dB HL. Audiogram, adaptive threshold measurements and just audible levels are within a range of 10 dB. For ACALOS, the levels at CU 5, 15, 25, and 35 were taken for the categories just audible, soft, comfortable, and loud, respectively. Almost no differences between the methods 1-3 can be observed. Levels of method 2 are slightly higher at 500 and 6000 Hz for the category loud. All three methods show lower levels for the categories comfortable and loud when compared to ACALOS. Standard errors are comparable for all three methods and ACALOS.

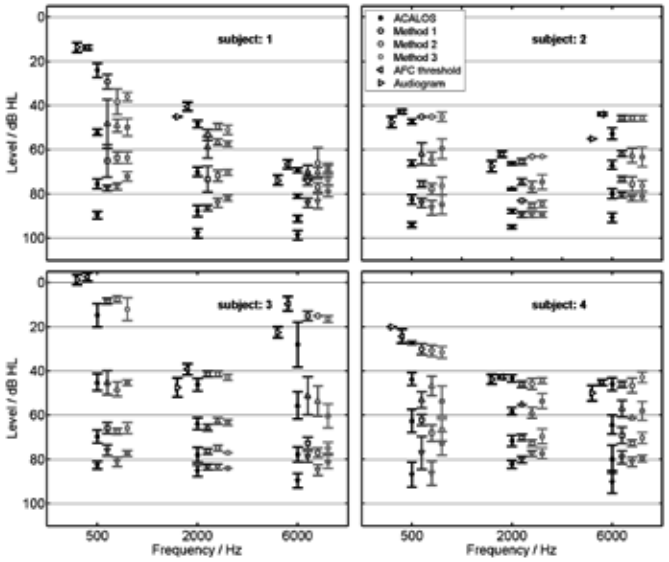

Fig. 3: Mean levels and standard deviation for 4 exemplary subjects. Subject 1 (male, 71 years) and 4 (female, 71 years) have an age-related sloping hearing loss at high frequencies whereas subjects 2 (male, 25 years) and 3 (female, 33 years) have a u-shaped audiogram.

As an example, mean values and standard deviations of 4 subjects are shown in Fig. 3. Standard deviations of the 4 repetitions of the suggested methods are comparable to those of ACALOS. Results of subject 1 show large standard deviations in method 1 at 500 and 2000 Hz. At 6000 Hz, the remaining dynamic range from just audible to loud derived from ACALOS is about 30 dB (70 to 100 dB HL). The suggested methods show a remaining dynamic range of 15-20 dB. Subject 3 shows a particularly large deviation at 6000 Hz for soft levels for methods 1-3 and ACALOS.

Comparison to ACALOS

All subjects show considerably higher levels for the category loud derived from ACALOS when compared to the suggested methods 1-3. To quantify this effect, the mean values for all 4 loudness categories for each subject were calculated. Mean levels for each loudness category across all 3 methods were taken to derive the corresponding CU values from the individual loudness function of each listener. The median and interquartile ranges are shown in Fig. 4. The suggested methods lead to lower levels when compared to ACALOS. Values for loud signals correspond to a CU value of 35 in ACALOS. Using the suggested methods 1-3, a level corresponding to 26.5 CU in the individual loudness functions was adjusted for loud. Values for comfortable levels were approx. 20 CU compared to 25 CU in ACALOS. Mean CU values derived from the individual loudness function led very similar values for normal-hearing and hearing-impaired listeners.

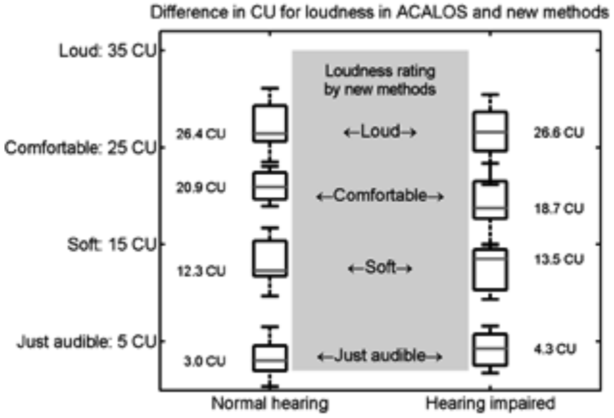

Fig. 4: To compare the suggested methods and ACALOS, the individual CU values for the 4 loudness categories were derived. The values are similar for normal-hearing and hearing-impaired listeners.

Towards hearing aid fittings

The results of methods 1-3 can serve as basis for loudness-based hearing aid or audio device fitting, restoring NH loudness perception for narrow-band noises (see red line in Fig. 1). NAL-NL2 was used as a reference method in a comparison of the following fitting methods:

Loudness based on ACALOS

In Latzel *et al.* (2004) the loudness-based fitting method is described, which restores the average narrow band loudness function measured with ACALOS in normal-hearing listeners. The difference between the loudness function of the hearing-impaired listener and normal-hearing listener leads to the desired gain (see Fig. 1).

Loudness based on method 1, 2, and 3

The same fitting rule as used for the ACALOS data was applied to the results of methods 1-3. The mean values of the normal-hearing listeners served as the reference values. The difference between the levels adjusted by the hearing-impaired listener and the reference values serve as the desired gain values. A linear, nearest-neighbor interpolation was used for levels between the adjusted categories.

Model based

The model-based fitting rule suggested by Ewert and Grimm (2011) uses the audiogram and the slope of the lower part of the ACALOS loudness function.

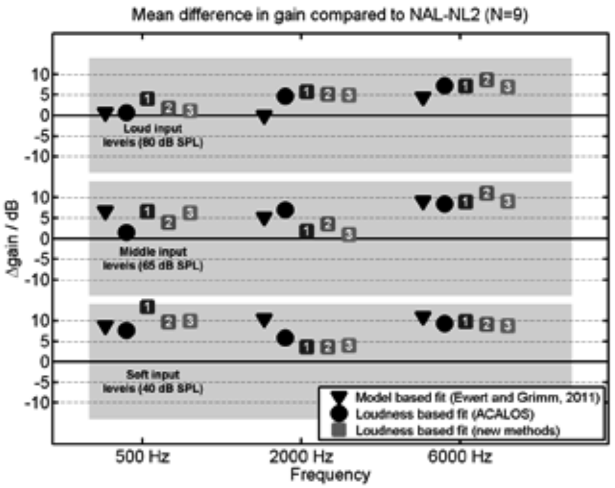

Fig. 5: Mean difference between the resulting gain values for each method and NAL-NL2 is shown. The loudness-based methods (ACALOS and methods 1-3) lead to very similar gain values, higher than the reference values from NAL-NL2.

Results of the gain calculation are shown in Fig. 5. Although very different procedures were used to measure the auditory capacity, they all led to similar gain values. This shows that the results of the suggested methods 1-3 are generally suited for measuring the hearing ability. As in case of loudness compensation based on ACALOS, the complete loudness compensation led to higher gains when compared to NAL-NL2.

SUMMARY AND DISCUSSION

User friendly, intuitive methods for measuring auditory capacity were developed and evaluated. The methods are fast: a single run for three frequencies was typically performed within 4-6 min including “just audible” loudness measurements in the new methods. Results are similar to ACALOS in terms of reproducibility but yield considerably lower levels for the categories comfortable and loud than ACALOS. This result appears plausible as the random nature of different levels presented in ACALOS tends to let listeners "save" the categories very loud and too loud for yet to come probably even louder sounds. In this way, the same levels lead to lower loudness categories as for the suggested adjustment and combined matching methods. Although the suggested methods are working in a quite different range of the individual loudness function, they led to almost the same gain prescription based on loudness compensation strategy for narrow band noises. The major advantage is that the suggested methods only use sound levels safely under the uncomfortable level which is a strong requirement for consumer audio devices. The higher gains compared to NAL-NL2 can lead to higher speech intelligibility as shown in Kreikemeier *et al.* (2011). However, higher gains might also lead to reduced acceptance for broadband stimuli. Currently, the inclusion of a loudness summation measure which would lead to reduced gain in such conditions is under investigation.

The three suggested methods showed almost identical results. Method 2 was approx. 60 s faster than method 1 and 3 and was often rated as the easiest method.

Further steps are the refinement of an appropriate fitting rule and integration into real devices (hearing aids, other audio devices) as well as evaluation in terms of loudness, quality, and speech reception.

[This work was supported by the BMBF ("Modellbasierte Hörsysteme").]

REFERENCES

ANSI (American National Standards Institute) (**2004**) "Specification for audiometers" (S3.6–2004), New York

Brand, T., and Hohmann, V. (**2002**). "An adaptive procedure for categorical loudness scaling" J. Acoust. Soc. Am. **112**, 1597-1604

Ewert, S.D., and Grimm, G. (**2011**). "Model-based hearing aid gain prescription rule" ISAAR 2011

Hougaard, S., and Ruf, S. (**2011**). "EuroTrak I: A Consumer Survey About Hearing Aids in Germany, France, and the UK" in *The Hearing Review*

Keidser, G., and Dillon, H. (**2006**). "What's New in Prescriptive Fittings Down Under?" in *Hearing Care for Adults*, pp. 133-142

Kreikemeier S., Latzel, M. and Kießling J. (**2011**). "Loudness-based Hearing Aid Fitting with Instantaneous Insitu-Percentile Monitoring" Z. Audiol. **50**, 62-72

Latzel, M., Margolf-Hackl, S., Denkert, J., and Kießling, J. (**2004**). "Präskriptive Hörgeräteanpassung auf Basis von NAL-NL1 im Vergleich mit einem lautheitsbasierten Verfahren" DGA.

Lecluyse, W., and Meddis, R. (**2009**). "A simple single-interval adaptive procedure for estimating thresholds in normal and impaired listeners" J. Acoust. Soc. Am. **126**, 2570-2579

Clinical measures of static and dynamic spectral-pattern discrimination in relationship to speech perception

STANLEY SHEFT, ROBERT RISLEY, AND VALERIY SHAFIRO

Department of Communication Disorders & Sciences, Rush University Medical Center, 600 South Paulina Street, suite 1012, Chicago, IL 60612 USA

Two experiments evaluated discrimination ability for both static and dynamic spectral patterns. The static conditions measured the ability to detect a change in the phase of a low-rate sinusoidal spectral ripple of wideband noise. The dynamic condition determined the signal-to-noise ratio (SNR) needed to discriminate 1-kHz pure tones frequency modulated by different 5-Hz lowpass noise samples drawn from the same underlying noise distribution so that discrimination was based on the temporal pattern of fluctuation. Both procedures used a modified descending method of limits with test stimuli recorded on a CD for clinic use. Results from the first experiment showed a significant relationship of both metrics to masked speech intelligibility. Using only the static procedure, the second experiment evaluated the role of fine-structure information in the perception of masked speech through vocoding of psychoacoustic and speech stimuli. In this case, results showed significant relationship only when the psychoacoustic and speech stimuli were either both vocoded or both unprocessed, consistent with involvement of stimulus fine structure in speech perception at low SNRs. Overall, results from both experiments support clinical utility of the procedures in the context of speech processing ability.

INTRODUCTION

Due to manner of production, speech can be represented by distinctive spectral patterns that vary over time. From this basis, past work has shown relationship between the ability to discriminate spectral patterns and measures of speech intelligibility in clinical subject groups. This work has evaluated auditory processing of both static and dynamic spectral patterns. A common approach in procedures that used static patterns was to assess the ability to either detect or discriminate periodic spectral rippling of wideband stimuli (*e.g.*, Litvak *et al.*, 2007; Won *et al.*, 2011). Past evaluation of dynamic spectral patterns has measured discrimination of either the rapid spectral variations of Schroeder-phase harmonic complexes or low-rate stochastic frequency modulation (FM) of pure-tone carriers (Drennan *et al.*, 2008; Sheft *et al.*, 2011). The current study represents initial efforts at developing clinically feasible measures of both static and dynamic spectral-pattern discrimination. Past work evaluating discrimination of static spectral patterns in clinical subject groups measured performance in terms of the threshold density of spectral rippling. So that density was constant at a value consistent with involvement

Proceedings of ISAAR 2011: Speech perception and auditory disorders. 3rd International Symposium on Auditory and Audiological Research. August 2011, Nyborg, Denmark. Edited by T. Dau, M. L. Jepsen, J. Cristensen-Dalsgaard, and T. Poulsen. ISBN 87-990013-3-0. EAN 9788799001330. The Danavox Jubilee Foundation, 2012.