SECOND-ORDER TEMPORAL MODULATION TRANSFER FUNCTIONS IN NORMAL-HEARING AND HEARING-IMPAIRED LISTENERS

Christian Lorenzi; Christian Füllgrabe; Jérome Sibellas; Frédéric Apoux
Institut Universitaire de France
Laboratoire de Psychologie Expérimentale
UMR CNRS 8581
Institut de Psychologie, Université René Descartes Paris 5
71 Av Edouard Vaillant, 92774 Boulogne-Billancourt Cédex, France
Tel: 33+ (0)1 55 22 57 34
Fax: 33+ (0)1 55 20 58 54
E-mail : christian.lorenzi@psycho.univ-paris5.fr

ABSTRACT

This paper reviews a series of studies that extend previous work on auditory time analysis in normal-hearing and hearing-impaired listeners. These studies are based upon the concept of "2nd-order" temporal modulation transfer functions (TMTFs) which are obtained by measuring detection thresholds for 2nd-order modulation (that is, sinusoidal modulation applied to the modulation depth of a sinusoidally amplitude-modulated tone or noise carrier), as a function of f_m', the rate of 2nd-order modulation. Here, the modulated tone or noise acts as a carrier stimulus of rate f_m and 2nd-order modulation is a complex envelope with 3 modulation components of rates f_m, f_m-f_m' and f_m+f_m'. Second-order TMTFs have been assessed in normal-hearing listeners and listeners with moderately severe cochlear damage. Overall, 2nd-order TMTFs are similar and lowpass in shape in both groups. The data suggest that 2nd-order TMTFs may be viewed as descriptions of the attenuation characteristics of the salient envelope beat produced by 2nd-order modulation at rate f_m'. The data also indicate that central auditory distortions may contribute to the perception of the temporal-envelope beat cue. This new paradigm is discussed in light of the linear systems analysis approach upon which the concept of (1st-order) TMTF was initially based.

Background

Speech conveys information both in its spectral and temporal structure. The temporal information in speech consists of fluctuations in amplitude at rates between about 2 Hz and 10 kHz. A subset of these temporal fluctuations - referred to as envelope fluctuations - correspond to the amplitude-modulation rates below about 50 Hz. Over the last decades, numerous studies on speech perception have shown that these slow variations in the temporal envelope can provide crucial cues for listeners with normal hearing [1,2] and listeners with cochlear hearing loss [3]. This has led to the suggestion that part of the reason people have difficulties understanding speech may be related to (i) degraded transmission of envelope cues [4], or (ii) abnormal ability to resolve changes in the sounds' envelope, that is, abnormal temporal acuity. Temporal-envelope sensitivity is traditionally assessed by measuring detection thresholds for sinusoidal amplitude modulation (SAM) as a function of modulation rate, f_m [5,6]. Inspired by the linear system analysis approach, this method attempts to describe auditory temporal acuity from the listeners' response to single-component temporal envelopes [5]. This approach has been applied to normal-hearing listeners [5,7], listeners with mild-to-severe cochlear hearing loss [6], cochlear implant users [8] and listeners with central auditory processing disorders [9,10]. In

normal-hearing listeners and patients with mild to severe cochlear damage, TMTFs resemble a lowpass filter with a cut-off frequency of either 50-60 Hz (when broadband noise is used as carrier) or 100-150 Hz (when a high-frequency pure tone is used as carrier). Moreover, in both groups, modulation detection thresholds are generally similar at most modulation rates, suggesting that the speech perception deficits shown by listeners with cochlear damage are not caused by abnormal auditory temporal acuity. By contrast, important variations in modulation sensitivity and TMTF shape have been reported for cochlear implant users and listeners with central auditory processing disorders. Cazals et al. (1994) [8] compared SAM detection and speech identification in a group of cochlear implant listeners. The results showed a significant correlation between TMTF slopes and consonant identification scores. Moreover, neuropsychological studies report a systematic relationship between speech intelligibility and envelope sensitivity at low modulation rates in listeners suffering from developmental dyslexia [10] and brain-damaged listeners with aphasia [9].

The linear systems analysis upon which the TMTF approach is based [5] has proved to be a powerful description of a number of aspects of auditory temporal processes and speech perception. Here, it is also important to note that some speech perception data based on temporal-envelope filtering techniques [1] rely heavily on the linearity of auditory temporal envelope processes. Evidence from recent electrophysiological and psychoacoustical studies [11,12] investigating the neural and perceptual response to two-component modulators may change this conception of auditory processes. In a linear system, there is no spectral energy at the envelope beat rate f_2-f_1 produced by two SAMs of rate f_1 and f_2. However, both studies revealed that nonlinear mechanisms in the auditory pathway generate an audible distortion component at the rate of the envelope beat produced by this two-component modulator. These data suggest that, despite its initial success, the linear systems analysis upon which the TMTF concept if based may be a limited approach. As a consequence, the numerous empirical data suggesting that the deficits in speech intelligibility shown by listeners with cochlear hearing loss are not due to impaired temporal resolution should be taken with caution: a different conclusion may be reached if sensitivity to the temporal envelope is measured with complex temporal envelopes (instead of SAM).
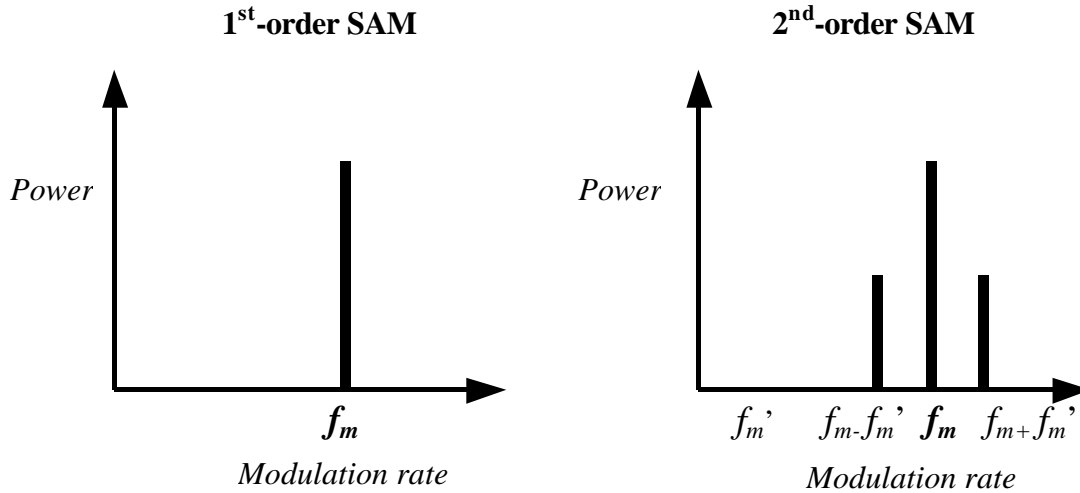

Figure 1: *Schematic modulation spectra of the 1st-order and 2nd-order SAM.*

In two recent studies (Lorenzi et al., 2001 [13,14]), detection thresholds were measured for sinusoidal modulation applied to the *modulation depth* of a sinusoidally amplitude-modulated noise or tone carrier as a function of the rate of the modulation applied to the modulation depth (referred to as f_m'). The sinusoidal modulation applied to the *modulation depth* of a sinusoidally amplitude-modulated carrier was referred to as "2nd-order" modulation (i.e., modulation *of* modulation) by the authors. In such stimuli, the first SAM acted as a "carrier" of rate f_m in the modulation domain, and sinusoidal modulation of the SAM modulation depth generated two additional components in the modulation spectrum at rates f_m-f_m' and f_m+f_m': the compound temporal envelope consisted therefore of the sum of three single SAM components. This is illustrated in Figure 1. The cyclic variation in modulation depth also produced a beat in the

stimulus' temporal envelope at a slow rate equal to f_m'; however, it is important to note that there is no energy in the modulation spectrum of the physical stimulus at the envelope beat rate f_m'. For each "carrier" modulation rate f_m, a "2nd-order" TMTF can be obtained by measuring the modulation depth, m', of the sinusoidal variation applied to the "carrier" modulation depth necessary to just detect the modulation, as a function of f_m'. Second-order TMTFs assess therefore the auditory sensitivity to complex temporal envelopes (instead of single component envelopes, as in the "classical" TMTF paradigm), and provide a new estimate of auditory temporal acuity that may highlight the contribution of peripheral and central nonlinearities to the perception of complex sounds such as speech. This paper reviews three studies that assessed 2nd-order TMTFs in normal-hearing and hearing-impaired listeners.

Second-order TMTFs in normal-hearing listeners

First and 2nd-order TMTFs were measured in 7 normal-hearing adult listeners for different carrier conditions. TMTFs were measured with white noise carriers in 4 listeners [13], and with sine carriers in 3 listeners [14]. The reader is referred to these publications [13,14] for details of the stimulus generation and procedures used in the detection tasks. Mean data across listeners are shown in the left and right panels of Figure 2 for the 5-kHz and white noise carriers, respectively. Consistent with previous studies [5,6,7], 1st-order TMTFs display a typical lowpass characteristic: Sensitivity is reduced by about 3 dB at f_m =64 Hz when the carrier is a white noise. The TMTF's cutoff frequency is higher when a 5-kHz pure-tone carrier is used. On average, SAM-detection thresholds increase slightly but continuously up to f_m =128 Hz, and sensitivity is reduced by 9 dB at f_m =256 Hz. For each "carrier" modulation rate f_m, 2nd-order TMTFs measured with the 5kHz pure-tone and noise carriers also display a lowpass shape: sensitivity to 2nd-order SAM degrades when f_m' increases.
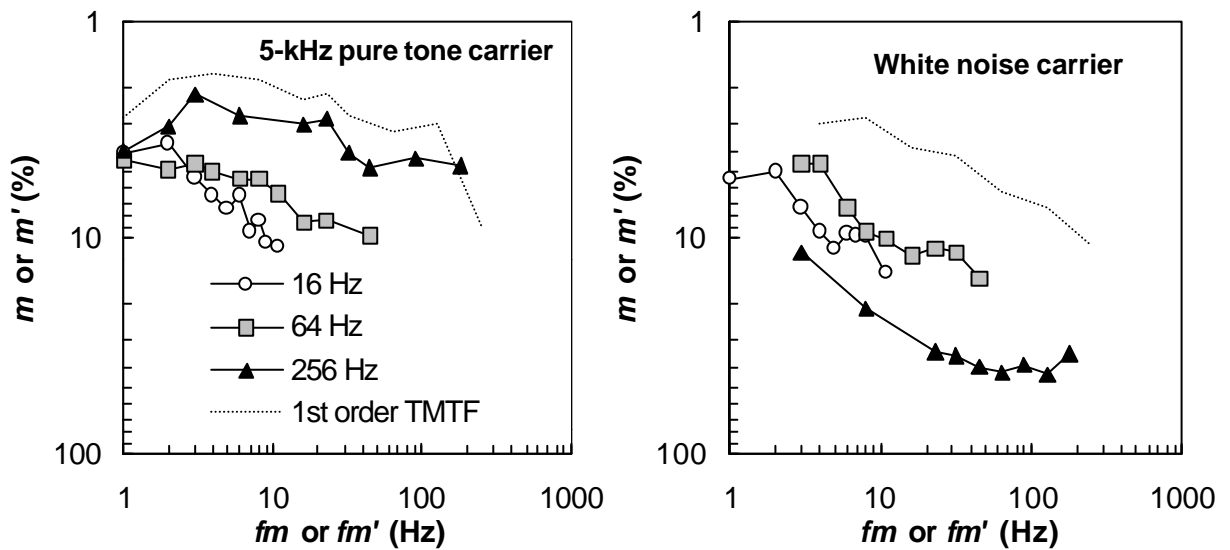

Figure 2. Left panel : Mean 2nd-order TMTFs for a 5kHz pure-tone carrier [14]. Right panel : Mean 2nd-order TMTFs for a white noise carrier [13]. In each panel, the 2nd-order TMTFs measured for f_m = 16 Hz, 64 Hz, and 256 Hz are plotted along with the 1st-order TMTF. In the case of 2nd-order TMTFs, the ordinate indicates 2nd-order modulation depth at threshold m', and the abscissa represents f_m'; in the case of 1st-order TMTFs, the ordinate indicates 1st-order modulation depth at threshold m, and the abscissa represents f_m.

These results indicate that auditory sensitivity to regular fluctuations in the strength of SAM is best for low fluctuation rates f_m'. For pure-tone carriers, overall sensitivity and cutoff frequency estimated from 2nd-order TMTFs increase when the "carrier" modulation rate f_m increases from 16 to 256 Hz. At the highest "carrier" modulation rate tested (256 Hz), 2nd-order modulation detection thresholds are very close to 1st-order modulation detection thresholds, indicating that 2nd-order SAM may be perceptually as salient as 1st-order SAM. For broadband noise carriers, overall sensitivity is best below about 2-4 Hz. It increases when the "carrier" modulation rate f_m

increases from 16 to 64 Hz, but degrades considerably at 256 Hz. In agreement with the results of the study by Dau et al. (1997 [15]) on 1st-order SAM detection, these results suggest that, when using a white noise carrier, (i) the intrinsic statistical fluctuations of this noise carrier mask the "carrier" and sidebands components of the 2nd-order SAM, and (ii) the masking effect is maximum at the highest modulation rate tested, i.e. 256 Hz.

An additional set of studies [14] investigated the role of envelope beat cues in 2nd-order SAM detection. Second-order SAM detection thresholds were measured with a 2-Hz wide narrowband noise carrier centred at 5 kHz, and compared with 2nd-order SAM detection thresholds measured with a 5-kHz sine carrier. The inherent statistical fluctuations of the narrowband noise carrier masked any component in the low frequency region ($\leqslant 2$ Hz) of the stimuli's modulation spectrum [15]. The mean results obtained on 3 listeners for $f_m = 256$ Hz are presented in Figure 3. They show that the masking effect produced by the statistical fluctuations of the noise was restricted to the lowest 2nd-order SAM rates ($f_m' \leq 16$ Hz). Similar results were obtained for $f_m = 16$ and 64 Hz. This reveals that detection of spectral energy at the envelope beat rate f_m' contributes to 2nd-order SAM detection. Nevertheless, the slow fluctuations of the noise did not abolish the ability to detect 2nd-order SAM, suggesting that additional envelope cues are involved in 2nd-order SAM detection.
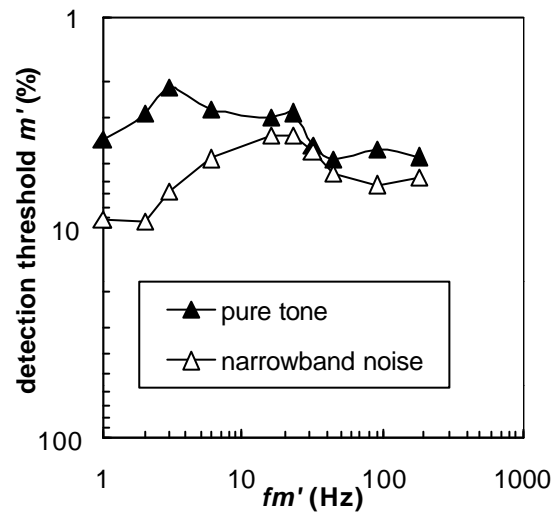

Figure 3. *Mean 2nd-order TMTFs measured with a 5-kHz pure-tone carrier and a 2Hz wide narrowband noise carrier centred at 5kHz [14].*

Second-order TMTFs in hearing-impaired listeners

The modulation masking data presented above emphasize the idea that nonlinear mechanisms in the auditory pathway produce audible effects in the temporal envelope domain [11,12]. The auditory nonlinearity may correspond to the fast acting compression performed by active mechanisms within the cochlea. Cochlear damage results in a loss of this compressive nonlinearity, and it should therefore lead to the abolition of the distortion component at the envelope rate f_m' produced by 2nd-order modulation *if* cochlear compression plays a critical role in the generation of such a distortion component. On the other hand, numerous studies have revealed that 1st-order modulation detection thresholds are generally unaffected by cochlear damage if audibility is controlled for [6]. Thus, cochlear damage should yield preserved 1st-order modulation detection thresholds and increased 2nd-order modulation detection thresholds. This hypothesis was tested by Tandetnik et al. (2001 [16]). To address this issue, 1st- and 2nd-order modulation detection thresholds were measured at low modulation rates for a 2s white noise carrier in listeners with sensorineural hearing loss and normal hearing listeners. In each task, listeners were tested at a comfortable listening level, ,which corresponded to 75 dB SPL for the normal-hearing listeners, and 90-100 dB SPL for the hearing-impaired listeners. Four normal-hearing listeners and 4 listeners with moderately-severe, bilateral symmetrical cochlear hearing loss and speech perception deficits participated in the experiments. Two impaired listeners showed a nearly flat hearing loss, while the two other impaired listeners showed a high-

frequency hearing loss. Detection thresholds for 1st and 2nd-order SAM were obtained using an identical 2I, 2AFC psychophysical procedure with feedback to that presented above [13,14]. The mean data are presented in Figure 4.
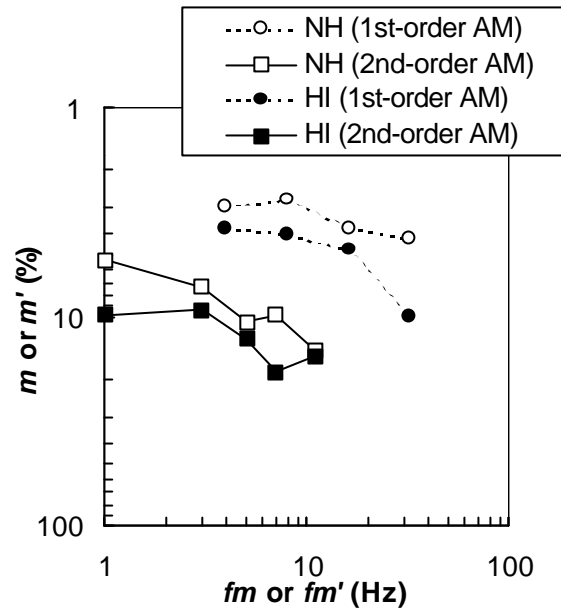

Figure 4: *Mean data for the normal-hearing (NH) and hearing-impaired (HI) listeners; 2nd-order modulation detection thresholds were measured for $f_m = 16$ Hz.*

The results showed that, in hearing-impaired listeners: (i) 1st-order modulation detection thresholds were within the normal range up to $f_m=16$ Hz, and poorer than normal at $f_m=32$ Hz, (ii) 2nd-order modulation detection thresholds were within the normal range at $f_m' = 3$, 5 and 11 Hz, and poorer than normal at $f_m' = 1$ and 7 Hz. When observed, the differences between normal-hearing and hearing-impaired listeners were relatively small (5 dB). Therefore, it appears that the loss of the peripheral compressive nonlinearity caused by cochlear damage only slightly degrades 2nd-order modulation detection. These results suggest that: (i) cochlear damage has little effect on the detection of both sinusoidal and complex temporal envelopes, and (ii) the nonlinear mechanisms producing the distortion component at the envelope beat rate should probably have a central instead of a cochlear origin.

Conclusions

The data presented above suggest that 2nd-order TMTFs may be viewed as quantitative descriptions of the attenuation characteristics of the envelope beat cue produced by complex temporal envelopes. Consistent with previous models of temporal envelope perception assuming that fast amplitude fluctuations are smoothed by the auditory system [5], slow envelope beats are better detected than fast ones. The 2nd-order SAM detection data also indicate that, under certain circumstances, envelope beats may be as salient as the single modulations composing any complex envelope. This emphasises therefore the potential contribution of such envelope beats to complex sound perception such as speech perception.

Our masking data [14] also indicate that the detection of this envelope beat is partly based on the detection of spectral energy at the envelope beat rate. This is inconsistent with the classic models of temporal acuity [5]. It is however compatible with an alternative model - the modulation filterbank model [15] - which assumes that the temporal envelope of sounds is decomposed by an array of broadly tuned filters, provided that auditory nonlinearities distort the internal representation of complex envelopes. Our data collected in hearing-impaired listeners suggest that the nonlinear mechanisms producing the distortion component should have a central (i.e., retrocochlear) instead of a cochlear origin. In contrast with our initial expectations, this also indicates that the speech perception deficits shown by hearing-impaired listeners are not due to impaired ability to resolve changes in the complex envelopes of speech sounds.

Finally, masking data suggest that additional envelope cues might be involved in 2nd-order SAM perception. Recent experiments conducted by Füllgrabe and Lorenzi [17] indicate that these additional cues may correspond to envelope beat cues appearing at the output of modulation filters tuned to the components of complex envelopes.

Taken together, the results presented above suggest that 2nd-order TMTFs provide a new framework for the systematic study of auditory time-analysis. By revealing (i) the perceptual salience and the importance of temporal envelope beats in complex-envelope perception, and (ii) the existence of auditory distortions in the internal representation of complex envelopes, the 2nd-order TMTF approach refines the initial framework upon which the 1st-order TMTF approach was based.

Bibliographical references

[1] Drullman, R., Festen, J. M., and Plomp, R. (**1994**). "Effect of reducing slow temporal modulations on speech reception", J. Acoust. Soc. Am. **95**, 2670-2680.
[2] Shannon, R. V., Zeng, F., Kamath, V., Wygonski, J., and Ekelid, M. (**1995**). "Speech recognition with primarily temporal cues," Science **270**, 303-304.
[3] Apoux, F., Crouzet, O., and Lorenzi, C. (**2001**). "Temporal envelope expansion of speech in noise for normal-hearing and hearing-impaired listeners: Effects on identification performance and response times," Hear. Res. **153**, 123-131.
[4] Houtgast, T., and Steeneken, H. J. M. (**1985**). "A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria," J. Acoust. Soc. Am. **77**, 1069-1077.
[5] Viemeister, N. F. (**1979**). "Temporal modulation transfer functions based upon modulation thresholds," J. Acoust. Soc. Am. **66**, 1364-1380.
[6] Bacon, S. P., and Viemeister, N. F. (**1985**). "Temporal modulation transfer functions in normal-hearing and hearing-impaired subjects," Audiology **24**, 117-134.
[7] Kohlrausch, A., Fassel, R., and Dau, T. (**2000**). "The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers," J. Acoust. Soc. Am. **108**, 723-734.
[8] Cazals Y, Pelizzone M, Saudan O, and Boex C. (**1994**). "Low-pass filtering in amplitude-modulation detection associated with vowel and consonant identification in subjects with cochlear implants," J. Acoust. Soc. Am. **96**, 2048-2054.
[9] Hescot F, Lorenzi C, Debruille X, and Camus JF. (**2000**). "Amplitude-modulation detection for broadband noise in a single listener with left-hemisphere damage," Brit. J. Audiol. **34**, 341-351.
[10] Lorenzi, C., Dumont, A., and Füllgrabe, C. (**2000**). "Use of temporal envelope cues by developmental dyslexics," J. Speech Lang. Hear. Res. **43**, 1367-1379.
[11] Shofner, W. P., Sheft, S., and Guzman, S. J. (**1996**). "Responses of ventral cochlear nucleus units in the chinchilla to amplitude modulation by low-frequency, two-tone complexes," J. Acoust. Soc. Am. **99**, 3592-3605.
[12] Moore, B. C. J., Sek, A., and Glasberg, B. R. (**1999**). "Modulation masking produced by beating modulators," J. Acoust. Soc. Am. **106**, 908-918.
[13] Lorenzi, C., Soares, C., and Vonner, T. (**2001**). "Second-order temporal modulation transfer functions, " J. Acoust. Soc. Am. **110**, 1030-1038.
[14] Lorenzi, C., Simpson, M. I. G., Millman, R. E., Griffiths, T. D., Woods, W. P., Rees, A., and Green, G. G. R. (**2001b**). "Second-order modulation detection thresholds for pure-tone and narrow-band noise carriers," J. Acoust. Soc. Am. **110**, 2470-2478.
[15] Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997**). "Modeling auditory processing of amplitude modulation : I. Modulation detection and masking with narrow-band carriers," J. Acoust. Soc. Am. **102**, 2892-2905.
[16] Tandetnik, S., Garnier, S., and Lorenzi, C. (**2001**). "Measurement of first and second-order modulation detection thresholds in listeners with cochlear hearing loss," Brit. J. Audiol. **35,** 355-364.
[17] Füllgrabe, C. and Lorenzi, C. (**2002**). "Detection and discrimination of 2nd-order SAM," 143rd Meeting of the Acoustical Society of America, Pennsylvania, 3-7 June 2002.