Title: Isolating Spectral Cues in Amplitude and Quasi-Frequency Modulation Discrimination by
Reducing Stimulus Duration

Article Type: Research Paper

Authors: Ewa Borucki, Ph.D.; Bruce G Berg, Ph.D.

Corresponding Author: Ewa Borucki— eborucki@uci.edu

Department of Cognitive Sciences

2201 Social & Behavioral Sciences Gateway Building (SBSG)

University of California

Irvine, CA 92697-5100 ** (same for both authors)

Abstract

This study investigated the psychophysical effects of distortion products in a listening task traditionally used to estimate the bandwidth of phase sensitivity. For a 2000 Hz carrier, estimates of modulation depth necessary to discriminate amplitude modulated (AM) tones and quasi-frequency modulated (QFM) were measured in a two interval forced choice task as a function modulation frequency. Temporal modulation transfer functions were often non-monotonic at modulation frequencies above 300 Hz. This was likely to be due to a spectral cue arising from the interaction of auditory distortion products and the lower sideband of the stimulus complex. When the stimulus duration was decreased from 200 ms to 20 ms, thresholds for low-frequency modulators rose to near-chance levels, whereas thresholds in the region of non-monotonicities were less affected. The decrease in stimulus duration appears to hinder the listener's ability to use temporal cues in order to discriminate between AM and QFM, whereas spectral information derived from distortion product cues appears more resilient.

Keywords: Amplitude Modulation; distortion products; cubic distortion tone; duration, phase sensitivity; bandwidth

I. INTRODUCTION

The primary aim of this study is to validate a new psychophysical technique that uses a discrimination task to investigate the perceptual effects of auditory distortion products. The heyday of psychophysical investigations of distortion product effects began with Goldstein (1967a) and continued throughout the 1970s, predating Kemp's (1978) discovery of otoacoustic emissions. The primary empirical method of that period was the probe-tone, beat-cancellation technique in which a probe-tone was added a few Hz away from the expected frequency of a distortion product in order to create the perception of beats. Listeners adjusted the amplitude and phase of a "cancellation tone" with the same frequency as the distortion product in order to cancel the perception of beats. The distortion product was assumed to have the same amplitude as the cancellation tone and a phase difference of π radians. Many of the basic discoveries of this period, such as the dominance of the difference tone (DT) and cubic distortion tone (CDT), were later confirmed with audiometric measurements of distortion product otoacoustic emissions (DPOAEs). Since that time, rather than a subject of psychophysical investigation, distortion product effects are often viewed as something to avoid, mask, or disregard with respect to both psychoacoustic experimentation and theory. For instance, the discrimination of amplitude modulated tones (AM) and quasi-frequency modulated tones (QFM), an ostensibly pristine technique for estimating phase sensitivity, is thought to be "contaminated" by distortion products (Goldstein 1967a; Buunen 1975).

In theory, an AM/QFM discrimination task is ideal for evaluating phase sensitivity as each stimulus has the same power spectra composed of three tones but having different envelopes as a result of a relative phase difference of the center tone. The center tone (i.e. carrier frequency) of the QFM stimulus is presented $\pi/2$ radians out of phase with the two sideband tones, whereas for AM, all three tones have the same phase. Consequently, QFM has an envelope modulation rate twice that of AM and a lower modulation depth. Mathes and Miller (1947) found that the two stimuli are audibly different

from one another when the modulation rate to center frequency ratio is less than 0.4. They posited that discrimination should only be possible when all three tones lie within the range of an auditory filter. Distortion products, however, were suspected of contributing to the discrimination of AM/QFM tones (Zwicker, 1955; Goldstein, 1967a). In particular, the interaction between the low sideband and an internally generated cubic distortion tone (CDT) produced by the two higher tones were thought to generate a spectral cue (i.e. intensity of the lower sideband) that aided in discrimination (Buunen, 1975; Goldstein, 1967b). The CDTs generated by AM and QFM tones, designated CDT_{AM} and CDT_{QFM} , have different phases and thus interact differently with the lower sideband. Given that a CDT is always present at same frequency as the lower sideband and that it is the most audible distortion product (Plomp, 1965), a CDT effect appears to be the most likely candidate of potential distortion product effects.

Although clear evidence for distortion product effects in AM/QFM discrimination did not exist at the time, few studies were completed after the hypothesis was articulated by Buunen (1975). Bernstein and Oxenham (2006) added low-pass noise in order to mask low frequency distortion products, but there may have been no ideal cutoff frequency for such a technique as the CDT interacts with the lowest tone of the stimulus. Without masking the CDT, it is unlikely that all pertinent distortion product effects were controlled. Increasing the cutoff frequency of the noise in order to mask the CDT would have also masked the lower sideband, thereby affecting the temporal envelope.

Nelson (1994), however, claimed that auditory distortion products are too weak to produce an intensity cue in AM/QFM discrimination. Nelson measured percent correct for AM/QFM discrimination and found that at higher intensities maximal bandwidth for phase sensitivity is broader than for lower intensity stimuli. However, Nelson's data exhibit some irregularities. In some cases, at lower intensity levels, percent correct improved at wider bandwidths after chance performance was obtained at narrower bandwidths. Irregularities found in Nelson's data diminish at higher intensities,

resulting in a higher percent correct and a wider estimated critical bandwidth for phase sensitivity. Nelson concluded that distortion products may induce envelope differences, but the magnitude of the CDTs themselves should not provide enough energy to induce a level cue; however, no thorough evaluation for that claim was given.

Evidence for a distortion product effect in AM/QFM discrimination was recently reported by Tabuchi, Borucki, and Berg (2012). They observed large non-monotonicities in temporal modulation transfer functions (TMTFs) (Viemeister, 1979) at high modulation rates, presumably attributable to the effects of the CDT. For a 2000 Hz carrier, for example, distinct minima in TMTFs were found at modulation rates greater than 300 Hz with thresholds that were as much as 35 dB lower than chance and 10 to 15 dB below thresholds obtained at lower modulation rates (e.g. 50 Hz). In a few cases of interest, referred to as “double-dip non-monotonicities”, the TMTF displayed two distinct minima of extreme sensitivity at high modulation rates with a small range of modulation frequencies in between where there was no measurable sensitivity. A “rotating-quadrant model” (RQM) was proposed that provided an explanation for non-monotonicities in TMTFs.

According to the RQM (Tabuchi, *et al.*, 2012) the interaction between the CDT and lower side band, designated f_L , introduces an intensity cue. The model draws on two basic findings from probe-tone, beat-cancellation experiments. First, the relative phase difference between CDT_{AM} and CDT_{QFM} is always $\pi/2$ radians, and second, the absolute phase of each advances as the frequency separation between the carrier (f_C) and upper side band (f_H) increases (Hall, 1972a,b). In terms of vector addition, it is useful to visualize two equal-length vectors representing the CDTs forming a quadrant positioned “on top” of a longer vector representing f_L . The distance from each respective “tip” of the quadrant to the base of f_L represents the combined amplitude of the CDT and lower sideband. As the stimulus bandwidth is increased, the “CDT-quadrant” rotates so that in some positions the tip-to-base distances are similar and there is no intensity cue, whereas in other positions, one tip of the quadrant is closer to

the base, representing an intensity difference. A complete rotation of the quadrant with increasing modulation rate produces a double-dip non-monotonicity in a TMTF that is qualitatively similar to data reported by Tabuchi, *et al.* They found that when the phase of the carrier is sampled from a uniform distribution with a modest range (e.g. $-\pi/3$ to $\pi/3$) on each stimulus presentation, thresholds at high modulation frequencies (i.e. the non-monotonicities in TMTFs) are degraded to chance or near-chance levels, whereas only a slight effect is found at lower modulation rates, implying that temporal cues are less affected by the phase randomization than spectral intensity cues. Moreover, reducing the range of phase randomization partially restores sensitivity at high modulation rates, a finding consistent with the RQM.

This study contributes complementary evidence that non-monotonicities in TMTFs are attributable to CDT effects. Whereas Tabuchi *et al.* (2012) degraded spectral cues while preserving temporal cues, the approach taken here is to degrade temporal information while preserving spectral cues by reducing the stimulus duration. Reducing duration should reduce listeners' abilities to discriminate between AM and QFM stimuli temporally, as they are given fewer looks at the stimulus envelope (Viemeister Wakefield, 1991). On the other hand, reducing duration should have less of an effect on a spectral intensity cue (Watson and Gengel, 1969) such that that the non-monotonicities in TMTF's are more resilient to the stimulus manipulation than thresholds at low modulation frequencies.

2 Method

2.1 Subjects

Seven subjects participated in the first experiment; all were affiliates of the University of California Irvine between the ages of 20 and 35. Listeners received monetary compensation for their participation. All subjects had pure-tone thresholds better than 20 dB HL for 8 kHz and below; middle

ear status was not evaluated.

2.2 Stimuli

AM and QFM stimuli are represented by:

$$y(t) = \sin(2\pi f_c t + \theta) + m/2 [\sin\{2\pi(f_c + f_m)t\}] + m/2 [\sin\{2\pi(f_c - f_m)t\}] \quad (1)$$

The stimulus defined by Eq.(1) is composed of a carrier frequency, f_c , and two sidebands, $f_L = f_c - f_m$ and $f_H = f_c + f_m$, where f_m is the modulation frequency of the AM stimulus and m represents the modulation depth of the stimulus. The stimulus $y(t)$ is defined as AM when phase, θ , is 0 and QFM when $\theta = \pi/2$.

Thresholds for modulation depth, m , are measured with a modified procedure where m is allowed to vary between 0 and 1. m varies in an one-up two-down adaptive procedure (Levitt, 1971), in 2 dB steps, in order to estimate thresholds for modulation depth required to discriminate between AM and QFM. A ceiling of $m=1$ must be imposed in order to avoid over-modulation. When an incorrect response is obtained with $m=1$, m remains at 1 until two correct responses are obtained. Stimulus values after the first four reversal points for all trials were averaged and taken as a threshold (Klein, 2001).

2.3 General procedure

AM and QFM were presented in random order in a 2IFC procedure. Subjects indicated which interval they believed contained the AM stimulus by pressing 1 or 2 on a keyboard and were given feedback after every trial. Each trial consisted of two 200 ms-presentation intervals with an inter-stimulus interval of 500 ms. The stimulus was gated by a raised cosine ramp for a duration of 10 percent of the stimulus at onset and offset. The level of the carrier was calibrated at 70 dB SPL. Sounds were generated digitally converted to analog with an external converter (E-MU 0202 Audio/MIDI

interface) at a sampling rate of 44100 Hz and presented diotically over Sennheiser eh350 headphones to subjects sitting in a single-walled sound attenuating chamber.

The carrier frequency was 2000 Hz while modulation frequency, f_m , was varied between 25 and 800 Hz. Subjects typically completed 5 to 10 blocks of 50 trials and 4-5 conditions during a 2 hour session, with conditions presented to subjects in random order in sets of 5 blocks. For all subjects, thresholds for modulation depth necessary to discriminate between AM and QFM stimuli were measured as a function of modulation rate.

For four subjects, those with nonmonotonic functions, thresholds for AM/QFM discrimination subsequently were estimated for shortened duration stimulus presentations; duration was limited to 20 ms in most cases (for one subject a 50 ms condition was used).

3 RESULTS

A TMTF for each listener is shown for the 200 ms condition in Figure 1. Listeners are divided into two groups, those with an essentially monotonic TMTF (upper panel) and those who exhibit a non-monotonic TMTF (lower panel). The dotted line at -2.65 represents chance performance (see Tabuchi et al., 2012). Having two types of listeners here is not a surprising result as large individual differences in audiometric measurements of CDTs exist (Gorga et al., 1996), so it is not expected that all subjects would be able to use the CDT as a cue in AM/QFM discrimination. Accordingly, Tabuchi, et al. (2012) found that individual listeners display nonmonotonic TMTFs for some carrier frequencies but monotonic TMTFs at others. For those with monotonic TMTFs, thresholds tend to indicate chance performance for listeners at a modulation rate around 300 Hz.

###figure 1##

The nonmonotonic TMTFs in the lower panel show a similar decrease in sensitivity at low modulation rates, but instead of approaching chance performance, performance begins to improve as the modulation rate exceeds 300 Hz. The data show some modest deviation between subjects at this transition point. For example, for S5 this transition occurs at 250 Hz.

For the listeners exhibiting a non-monotonic TMTF, a local minimum is found at a modulation frequency of 400 Hz, with the exception of S7 whose minimum lies at 500 Hz. A modulation rate of 400 Hz corresponds to an f_2/f_1 ratio of 1.2 (i.e. $f_1= 2000$ Hz and $f_2= 2400$ Hz, where f_1 and f_2 represent f_C and f_H respectively) . The overall range of f_2/f_1 ratios represented in this study is 1.025 to 1.4. DPAOE studies have shown that DPOAEs for the CDT tend to reach a maximal magnitude at an f_2/f_1 ratio 1.22 (e.g., Harris *et al.*, 1989; Gaskill and Brown, 1990). However, DPAOE measurements of the DP come from multiple sources, the best understood being the reflector and generator emissions, where the generator emission is thought to come from near f_2 place and the reflector from reflections from random impedance variations around the site of the DP (Talmadge, *et al.*, 1998; Shera and Guinan, 1999). It appears generator sources dominate perception as emission studies have found to be generator-type emissions largest at ratios between 1.1 and 1.3 and the levels of these emissions have been found to be more consistent with local non-linearities than the reflector type (Knight and Kemp, 2001; Dong and Olson, 2010). The increase in thresholds at modulation rates greater than 400 Hz is likely due to two factors. First, according to the RQM (Tabuchi, *et. al*, 2012), the phases of CDT_{AM} and CDT_{QFM} will change to a non-detect state as the bandwidth increases. Second, the strength of the DPOAE emissions decreases as f_2/f_1 exceeds 1.3. The latter suggests a necessary modification of the RQM in which the lengths of the vectors representing the CDTs and vary dependent on f_2/f_1 .

##fig 2##

As shown in Fig. 2, decreasing the stimulus duration has the greatest effects at low modulation rates. Closed symbols represent the thresholds obtained with a stimulus duration of 20 ms, which is compared to thresholds for the 200 ms condition replotted from Fig. 1 as open symbols. For S4, data is also presented for a 50 ms duration, represented by the open triangles. This listener presents the clearest demonstration that the cues for discrimination are different for low and high modulation rates. When the modulation rate is less than 200 Hz, thresholds increase systematically as the stimulus duration is decreased. The segments of the TMTFs for rates less than 200 Hz shift in parallel for the three different durations to a level of chance performance. In comparison, reducing the duration to 50 ms has only a small effect on the non-monotonicity. The location of the minimum threshold shifts from 400 to 450 Hz with little change in the overall span and depth of the non-monotonicity. When the duration is further reduced to 20 ms, the span of the non-monotonicity decreases, but there is little effect on the minimum threshold. S5 withdrew from the study before completing the 20-ms condition, but the data are consistent with the results for S4. The single point for S5 at a modulation rate of 75 Hz is at chance performance, whereas the two points at high modulation rates show only a slight increase in thresholds, and it may be the case that this increase simply reflects a shift in the minimum, as observed for S4.

S6 and S7 exhibit measurable sensitivity at low modulation rates when the stimulus duration is reduced to 20 ms. For S6, the low frequency segment of the TMTF shows a roughly parallel upward shift similar to that found for S4, but S7 presents a different case in that the increases in thresholds at 50 Hz and 100 Hz are much greater than the slight increase found at 200 Hz. The result suggests that a CDT-induced spectral cue is still available to S7 at 200 Hz when the temporal information is degraded by decreasing the stimulus duration. Conversely, S7 also exhibits the greatest effects of reducing the stimulus duration at high modulation rates in absolute terms, but one must consider the exceptionally low minimum threshold of -40 dB for a modulation rate of 500 Hz, 13 dB lower than the threshold at 100 Hz. Both S7 and S6 continue to produce a robust non-monotonicity for the 20 ms duration.

4 DISCUSSION

AM/QFM discrimination using tonal carriers offers a unique way to study temporal sensitivity as well as the perceptual effects of distortion products. Theorists have long hypothesized that measurements of temporal sensitivity with an AM/QFM discrimination task may be compromised by cues induced by distortion products. It was not predicted that these effects would manifest as non-monotonic TMTFs, nor was it predicted that temporal and spectral cues would have different dominance regions with respect to modulation frequency. Tabuchi, *et al.* (2012) showed that thresholds at high modulation rates were degraded by randomizing the phase of the carrier, with little effect on thresholds at low modulation rates. It was shown here that thresholds at low modulation rates were degraded by decreasing the stimulus duration, with little effect on thresholds at high modulation rates. Perceptually, subjects reported listening to “pitch”, “heaviness”, or “loudness” differences at high modulation rates and “roughness” at low modulation rates.

The finding that temporal and spectral cues dominate at different regions of the TMTF increases the validity of estimates of temporal sensitivity. Although an overlapping region of saliency for both temporal and spectral cues might be substantial in some cases, as suggested by the data of S7 in Fig. 2, it can be concluded that the maximum bandwidth of phase sensitivity is in the range of 500 Hz to 600 Hz. A carrier frequency to bandwidth ratio of 3.3 is generally consistent with Tabuchi, *et al.* (2012) for a 2 kHz carrier and with Bernstein and Oxenham (2006) for a 1 kHz carrier, but greater than ratios of 2.2 at 60 dB and 1.20 at 80 dB reported by Nelson (1994). The wider estimates of phase sensitivity might be due to a difference in task, AM/QFM discrimination with 100% modulation depth, or to a difference in intensity, but it is notable that the psychometric functions for a number of Nelson’s listeners were non-monotonic, particularly at low intensities, suggesting the presence of a distortion product effect. Nelson concluded that distortion product induced differences in the envelopes of AM and QFM stimuli was the cause of the nonmonotonic data and claimed that the magnitude of CDTs could not provide

enough energy to induce a level cue. If this were the case in the present study, then even if differences in envelopes were attributable to distortion products, the envelope cue should still be degraded with decreasing stimulus duration. If the range of non-monotonicities in this study were to be misinterpreted and included as estimates of phase sensitivity, then the estimates would be comparable to those found by Nelson.

The presence of non-monotonicities at a $1.2 f_2 f_1$ ratio suggests tonal stimuli are a viable method for evaluating distortion products allowing for a future method of measurement to be devised at a perceptual level rather than DPOAE output for humans. Although in any harmonic complex, there are numerous DPs generated, the CDT is likely the dominant source of discriminable information. Although the $2f_2 - f_1$ distortion product also interacts with the stimulus, (f_C and f_L creating an interaction with f_H), it does not appear to contribute much to discrimination. As shown by Tabuchi, *et al.* (2012), randomizing the carrier frequency or phase of f_C has little effect at low modulation rates and can eliminate non-monotonicities present at higher modulation rates; if the $2f_2 - f_1$ DP had a noticeable impact on discriminating modulated tone complexes, its effect should be apparent at lower modulation frequencies as it tends to be more robust at lower frequency ratios than the DP $2f_1 - f_2$ (Fitzgerald and Prieve, 2005; Knight and Kemp, 2001). Moreover, it would be expected that there would be more variability across all subjects for modulation frequencies below 300 Hz due to nature of individual differences in DPs as described above. Further research is required as it certainly cannot be asserted that there is no effect of the $2f_2 - f_1$ DP. Other DP are generated by the stimuli such as $2f_L - f_H$ and $2f_L - f_C$; however, these at the sights of non-monotonicities lie outside of an auditory filter at 800 and 1200 Hz respectively for a stimulus centered at 2000 Hz. While they may be detected, they may not serve as direct discrimination cues in a 2IFC as they are unresolved components not interacting with the stimulus and further interactions would be small. The addition of low-pass noise would remove ambiguity around the contribution of such DPs; however, that was beyond the intent of this study

Considering the large effect sizes seen here and in Tabuchi, *et al.* (2012), evaluating distortion products at the perceptual level could have advantages to traditional DPOAE measurements. DPOAE measurements are subject to external factors such as stiffness of the ossicular chain or calibration and probe factors which can be difficult to normalize. Beyond external factors, DPOAE measurements are subject to internal complications as well. By measuring DPs at a perceptual level, it may not be necessary to unmix sources (Talmadge, *et al.*, 1998; Talmadge, *et al.*, 1999; Kalluri & Shera, 2001; Shera and Guinan 2008). It has also been posited that notches in DP-grams come from a variety of sources more broadly distributed on the basilar membrane (e.g. Martin *et al.*, 2013). By measuring DPs at a perceptual level, it may not be necessary to unmix sources; however, what effect stimulus duration has when shifting the tone complexes studied here higher or lower in frequency would help elucidate the potential role of the reflection and generator components on discrimination; as stimulus durations are reduced, the less available the stimulus to interact with the CDT. Given inherent complications with tasks which involve human judgment and the complex nature of distortion products, more research is needed to clarify the roles of DPs the discriminability of tone complexes based on phase differences alone .

Acknowledgements

The authors would like to thank Dr. Virginia Richards, Dr. Kourosh Saberi, Dr. Brenda Lonsbury-Martin, and reviewers for their helpful comments. This work was supported by a grant from the National Science Foundation (BCS-07464003)

References

Bernstein, J.G. and Oxenham, A.J. (2006). The relationship between frequency

281 selectivity and pitch discrimination: sensorineural hearing loss. J. Acoust. Soc. Am. 120, 3929-
 282 3945.

283 Buunen, T.J.F. (1975). Two hypothesis on monaural phase effects. Acustica **34**, 98-105.

284 Eddins, D.A. (1999). Amplitude-modulation detection at low- and high-audio
 285 frequencies. J. Acoust. Soc. Am., **105**, 829-837.

286 Fitzgerald, T. S. and Prieve, B.A. (2005). “Detection of Hearing Loss Using $2f_2-f_1$ and
 287 $2f_1-f_2$ Distortion-Product Otoacoustic Emissions” J. of Speech, Language, and Hearing Res.,
 288 **48**(5), 1165-1186.

289 Gaskill, S. A. and Brown, A. M. (1990). “The Behavior of the Acoustic Distortion
 290 Product, $2f_1-f_2$, from the Human Ear and its Relation to Auditory Sensitivity,” J. Acoust.
 291 Soc. Am. **88**, 821–839.

292 Goldstein, J.L. (1967a). Auditory spectral filtering and monaural phase perception. J.
 293 Acoust. Soc. Am. **41**, 458-479.

294 Goldstein, J.L. (1967b). Auditory nonlinearity. J. Acoust. Soc. Am. **41**, 676-689.

295 Hann, H.H., Pratt, S.R., and Durrant, J.D., Parameters to Maximize $2f_2-f_1$ Distortion
 296 Product Otoacoustic Emission Levels. J. Speech, Language and Hearing Res. **51**, 1620-1629.

297 Hall, J.L. (1972a). Auditory distortion products f_2-f_1 and $2f_1-f_2$. J. Acoust. Soc. Am. **51**,
 298 1963-1871.

299 Hall, J.L. (1972b). Monaural phase effect: Cancellation and reinforcement of distortion
 300 products f_2-f_1 and $2f_1-f_2$. J. Acoust. Soc. Am. **51**, 1972-1881.

301 Harris FP, Lonsbury-Martin B.L., Stagner B.B., Coats AC, Martin G.K. (1989) Acoustic
 302 distortion products in humans: Systematic changes in amplitude as a function of f_2/f_1 ratio. J
 303 Acoust. Soc. Am. **85**,220-229.

304 Kemp, D.T. (1978). Stimulated acoustic emissions from within the human auditory

305 system. J. Acoust. Soc. Am. **64**, 1386-1391.

306 Kalluri, R., & Shera, C. A. (2001). Distortion-product source unmixing: A test of the
 307 two-mechanism model for DPOAE generation. J. Acoust. Soc. Am. **109**, 622–637.

308 Klein, S.A. (2001). Measuring, Estimating and Understanding the Psychometric
 309 Function: A Commentary. Percept. Psychophys. **63**, 1421-1455.

310 Knight, R. D. , and Kemp, D. T. (2001). “Wave and place fixed maps of the human ear,”
 311 J. Acoust. Soc. Am. **109**, 1513–1525.

312 Levitt, H. (1971). Transformed up-down methods in psychophysics. J. Acoust. Soc. Am.
 313 **49**, 467-477.

314 Martin, G.K., Stagner, B.B., and Lonsbury-Martin, B.L. (2013). Time-domain
 315 demonstration of distributed distortion-product emission components. J. Acoust. Soc. Am. **134**,
 316 324-355.

317 Mathes, R.C. and Miller, R.L. (1947). Phase effects in monaural perception. J. Acoust.
 318 Soc. Am. **19**, 780-797.

319 Nelson, D.A. (1994). Level-dependent critical bandwidth for phase discrimination. J.
 320 Acoust. Soc. Am. **95**, 1514-1524.

321 Plomp, R. (1965). Detectability threshold for combination tones. J Acoust. Soc. Am. **37**,
 322 1110-1123.

323 Shera C.A. and Guinan J.J, Jr., (1999). Evoked otoacoustic emissions arise by two
 324 fundamentally different mechanisms: a taxonomy for mammalian OAEs. J Acoust Soc Am **105**,
 325 782–798.

326 Tabuchi, H., Borucki, E., Berg, B.G. (2012). Effects of randomizing phase on the
 327 discrimination between amplitude-modulated and quasi-frequency-modulated tones, Hear. Res.
 328 **290**, 72-82.

329 Talmadge C.L., Long G.R., Tubis A., Dhar S. (1999) Experimental confirmation of the
 330 two-source interference model for the fine structure of distortion product otoacoustic emissions.
 331 J Acoust Soc Am **105**:275–292.

332 Viemeister, N.F. (1979). Temporal modulation transfer functions based upon modulation
 333 thresholds. J. Acoust. Soc. Am. 66,

334 Viemeister, N.F., Wakefield G.H. (1991). Temporal integration and multiple looks. J.
 335 Acoust. Soc. Am. **90**, 868-865.

336 Watson, C.S., Gengel, R.W. (1969). Signal duration and signal frequency in relation to
 337 auditory sensitivity. J. Acoust. Soc. Am. **46**, 989-997.

338 Zwicker, E. (1955). Der ungewöhnliche amplitudeng and der nichtlinearen verzerrungen
 339 des Ohres, Acustica. **5**, 67-74.

340 Zwicker, E. (1981). Dependence of level and phase of the $(2f_1-f_2)$ -cancellation tone on
 341 frequency range, frequency difference, level of primaries, and subject. J. Acoust. Soc. Am. **70**,
 342 1277-1288.

343
 344