Hearing Research xxx (2017) 1–5

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research Paper

Isolating spectral cues in amplitude and quasi-frequency modulation discrimination by reducing stimulus duration

Ewa Borucki, Ph.D. *, Bruce G. Berg, Ph.D.

Department of Cognitive Sciences, 2201 Social & Behavioral Sciences Gateway Building (SBSG), University of California, Irvine, CA 92697-5100, USA

A R T I C L E I N F O

Article history:
Received 22 June 2016
Received in revised form
1 March 2017
Accepted 2 March 2017
Available online xxx

Keywords:
Amplitude modulation
Distortion products
Cubic distortion tone
Duration
Phase sensitivity
Bandwidth

A B S T R A C T

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.

1. 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 cancelation 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

* Corresponding author.
E-mail address: eborucki@uci.edu (E. Borucki).

http://dx.doi.org/10.1016/j.heares.2017.03.004
0378-5955/© 2017 Published by Elsevier B.V.

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 et al. (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 10–15 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 and 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 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, $f_L = f_C\text{-}f_m$ and $f_H = f_C\text{-}f_m$, 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-