Peripheral analysis of frequency in human ears revealed by tone burst evoked otoacoustic emissions

Li Xu, Rudolf Probst, Frances P. Harris *, Jutta Roede

Department of Otorhinolaryngology, Kantonsspital Basel, CH-4031 Basel, Switzerland

(Received 27 May 1993; Revision received 19 October 1993; Accepted 11 November 1993)

Abstract

Otoacoustic emissions were evoked in the same ears with single tone bursts at 1, 2 and 3 kHz and with a complex stimulus consisting of a digital addition of the three tone bursts. Stimuli were presented at 75, 59 and 37 dB SPL to 28 ears of human subjects with normal hearing. The purpose was to determine if comparisons of responses to the complex stimulus with a posthoc addition of responses from single tone bursts could delineate features of cochlear frequency analysis of short-duration signals. For processing of the data, the results from the individual tone bursts were combined offline to form a composite response. This was then compared with the response obtained with the complex stimulus. Results revealed close correspondence between the spectra of the complex and composite responses in all ears despite interindividual differences in response morphology. Correlations between the complex and composite waveforms exceeded 80% for all stimulus levels. Subtractions of the two spectra revealed that the majority of the differences occurred at frequencies on the high-frequency slopes of the 1- and 2-kHz spectral peaks. This was due to a reduction in energy for the responses obtained with the complex stimulus. There was little variation between the two response types in the peak frequencies of their spectra, in the energy at frequencies on the lower frequency sides of the spectral peaks at 1 and 2 kHz, or in the spectral components at 3 kHz. Results reveal characteristics of the analysis of frequency in the preneural stages of cochlear processing.

Key words: Otoacoustic emissions; Frequency analysis; Cochlear mechanics; Transiently evoked otoacoustic emissions; Inner ear; Cochlear frequency selectivity

1. Introduction

A primary function performed by the peripheral auditory system is the analysis of frequency. The final preneural stages of this processing in the normally functioning cochlea involve macro and micromechanical events that are related to the frequency-place map. A broadly tuned passive element combined with additional components that may or may not be 'active' contribute to the ultimate sharp tuning that is observed in the response of the basilar membrane in vivo (e.g. Rhode, 1971; Robles et al., 1986). Frequency analysis by the cochlea can be assessed using various physiological and psychoacoustic methods. Recently, the strong link between otoacoustic emissions (OAEs) and cochlear mechanical events has made it possible to use OAE measurements for making inferences about cochlear activity. This is particularly useful with human subjects because the method is noninvasive, objective and often has clinical applicability.

Results of investigations of OAEs evoked by transient stimulation of the human ear have determined that the responses are frequency specific and relate to frequency analysis. Comparisons of responses evoked by tone burst and click stimuli in the same ear are a means of deriving this evidence (Elberling et al., 1985; Harris and Probst, 1990; Norton and Neely, 1987; Probst et al., 1986; Stover and Norton, 1991). The spectrum of a tone burst evoked OAE has energy that is limited to frequencies in the stimulus (Elberling et al., 1985; Norton and Neely, 1987; Probst et al., 1986). Comparisons of the spectra from click-evoked and tone burst-evoked responses in the same ear show that energy peaks occur at similar frequencies regardless of the form of the stimulation. That is to say, if tone
bursts are used to evoke responses at frequency regions that are either high or low in amplitude in the click-evoked spectrum, then the energy in the respective responses tends to correspond. It is possible to derive one type of response from the other by either separating frequency components of a click-evoked response or combining components from tone burst-evoked responses in the same ear. Probst et al. (1986), using the latter method, observed a close correspondence between a composite spectrum derived from single tone burst responses and the spectrum of a click-evoked response. With this relatively crude comparison, they confirmed that TEOAEs possess the property of linear superposition, which had been previously determined in the time domain by Zwicker (1983). An example of a response for one ear obtained using a complex stimulus with peaks at 1 and 3.6 kHz and individual tone bursts at the same frequencies was published by Kemp, Ryan and Bray (1990). When responses to the individual tone bursts were combined and the result compared with the response to the complex stimulus, there was a close correspondence of the two spectra.

In performing comparisons between the combined and click-evoked spectra in their results, Probst et al. (1986) noted that in some ears there were peaks present in the tone burst response that were absent in the click-evoked response. This implied that additional processes could be operating to produce the final output from click stimuli that are not invoked when stimulation is limited to a narrower frequency range. The superposition property may not be strictly linear. However, the nature of the interactions of the components in the click-evoked response that would produce these differences were not further characterized in this investigation.

By exploiting the features of frequency specificity and linear superposition of OAEs, it should be possible to parse out the interactions that occur when the cochlea separates a complex signal into its component frequencies. However, a click may not be the ideal stimulus for such an investigation because it would be difficult to produce individual stimuli that, when digitally added would produce essentially the same stimulus as the click. By using stimuli that restrict the range of generation to different sites contributing to the OAE and comparing the emissions to a simultaneous presentation of the frequency-restricted stimuli, the characteristics of cochlear frequency analysis of complex signals could be determined more precisely.

We compared the otoacoustic emissions evoked in individual ears by three tone bursts presented either singly or combined as a three-tone burst complex to determine: (a) the linearity of TEOAE generators during the integration and separation of frequency; (b) the relation of these findings to cochlear frequency analysis.

2. Methods

Subjects
Twenty-eight subjects in good general health (15 men and 13 women) ranging in age from 19 to 35 years (26.6 ± 4.6 years) participated. Both ears of each subject were evaluated; however, results from only one ear were included in the data analyses. The ear satisfying the following criteria (in order of priority) was chosen for inclusion: (1) better hearing, (2) higher amplitude OAEs for click stimuli, (3) presence/absence of spontaneous otoacoustic emissions (SOAEs). There were 15 right and 13 left ears included. The pure-tone air conduction thresholds of these ears were ≤ 15 dB HL from 0.25 to 8 kHz and their middle ear status was normal by otoscopic and impedance examinations.

Fourteen ears had SOAEs, as measured with methods previously described (Lonsbury-Martin et al., 1990) and 14 ears did not. The number of SOAEs ranged from 1 to 9 per ear with an average of 5 and their frequencies ranged from 585 to 7062 Hz with a dominant range from 1 to 2 kHz. All of the 28 ears had transiently evoked OAE responses from click stimuli.

Instrumentation and stimuli
A commercially available instrument, ILO88 Otodynamic Analyzer linked with a Compaq Portable III computer, was used to test TEOAEs. Two types of stimuli were generated by customized routines of the ILO88 software (version 3.92): (a) Cosine-windowed tone bursts of 5-ms duration (rise-fall time = 2.5 ms, plateau = 0 ms) with center frequencies of 1, 2 and 3 kHz, (b) A complex stimulus consisting of a digital addition of the three single tone burst stimuli. Stimuli were presented every 20 ms. The levels of each of the four stimuli were approximately 75, 57 and 39 dB SPL as measured in the ear canal. At 57 and 39 dB SPL, all stimuli were presented at the same level and polarity (‘linear presentation mode’). For stimuli at 75-dB SPL, a ‘nonlinear presentation mode’ was adopted. In this mode, a series of four stimuli were delivered with three at the same level and polarity and a fourth three times greater in level and inverted in polarity. This technique was selected for stimuli at the highest level to cancel the linear portions of the stimuli and the response (Kemp et al., 1986; Bray and Kemp, 1987). It was determined by testing the system in various passive cavities that stimulus artifacts did not influence the lower level results and the nonlinear technique was not necessary. Additionally, for the 75-dB stimuli, the spectrum of the tone burst complex was compared with a spectrum resulting from the digital sum of the three individual tone bursts. The spectral components of the stimuli were almost identical except for small differences in the sideband regions of the single tone bursts. An analysis window of 5.5–20.5 ms after the onset of
the stimuli was used to evaluate the responses, which resulted from 260 averages.

Procedure

Subjects were seated in a sound-treated room during the testing, which lasted about 30 min for each ear. Transiently evoked otoacoustic emissions were recorded with the three individual tone burst stimuli and the complex stimulus at each of the three levels. The 12 stimulus conditions (three tone bursts and the complex stimulus at each of three levels) were presented randomly to avoid order effects. Data were stored on floppy disk for analysis off line.

Analysis

Results were analyzed in both the frequency and time domains. In the frequency domain, the spectra of the individual responses obtained with all stimuli were processed in three stages: (1) Response and noise levels below 0.4 kHz and above 5 kHz were removed because no energy was present in the responses beyond these frequencies; (2) The noise levels, which varied slightly by frequency, were equalized to $M + 1.3$ SD based upon the mean noise level in the response from the complex stimulus at each level for each ear; (3) The spectra of the individual responses to tone bursts at each level were combined offline to form a composite response. Following these steps, the composite and complex responses were compared.

In the time domain, the individual waveforms generated by single tone bursts at the same stimulus level were added together using a software option available with the IL088. This resulted in a composite waveform that was compared with the waveform resulting from the complex stimulus. Correlational analyses of the two waveforms were then performed, which produced measures of the overall correlation in percentage and within 1-kHz bands centered at 1, 2 and 3 kHz.

The presence or absence of a response by frequency was determined based on these criteria: (1) The response contained at least three consecutive datapoints at frequencies corresponding to the stimulus region; (2) At least one of these datapoints was greater than or equal to 5 dB above the noise floor; (3) Review of the time-domain waveforms for responses that were questionable.

Table 1 lists the number of ears with responses at each stimulus level and for each single tone burst level.

![Fig. 1. Example of the correspondence of the spectra from the composite and complex responses from one ear with spontaneous otoacoustic emissions (SOAEs). Leftmost panels are the spectra for stimulus levels of either 75, 57 or 39 dB SPL. The horizontal dashed lines at the base of the traces represent the normalized noise floor for the responses. The rightmost panels are the result of the subtraction of the two spectra (Composite – Complex). The arrowheads along the abscissa indicate the frequencies of the SOAEs. There was good correspondence of the two spectra for this ear despite the complex configuration of the responses and the presence of SOAEs.](image-url)
Table 1

<table>
<thead>
<tr>
<th>toneburst frequency</th>
<th>1 kHz</th>
<th>2 kHz</th>
<th>3 kHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOAE/ + -</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stimulus level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>75 dB SPL</td>
<td>14</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>57 dB SPL</td>
<td>14</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>39 dB SPL</td>
<td>12</td>
<td>8</td>
<td>11</td>
</tr>
</tbody>
</table>

Total number of ears with SOAEs: 14
Total number of ears without SOAEs: 14

The highest level of stimulation elicited responses in all of the 28 ears except at 3 kHz for one ear without SOAEs. Stimulation at 57 dB SPL failed to elicit responses at 3 kHz in two ears without SOAEs. Decreasing the stimulus level to 39 dB SPL resulted in an overall reduction in the number of responses for each of the tone bursts. At this stimulus level, more ears with SOAEs retained responses than did those without SOAEs.

Statistical analyses were performed using Statview II software and included simple correlation and analysis of variance for repeated measures.

3. Results

Comparison of the composite and complex responses revealed a very close correspondence within all ears despite interindividual differences in response morphology and amplitude. The leftmost panels of Figs. 1 and 2 provide examples of the spectra from two ears, one with relatively good correspondence (Fig. 1) and the other with poorer correspondence (Fig. 2) between the composite and the complex responses. The noise floors have been equalized as described in the Methods section. The solid lines represent the composite spectra of the responses to the three tone bursts presented separately and the thicker dotted lines the spectra of the responses to the complex stimulus. Peaks in the responses always occurred at the same frequencies whether the stimulus was a single-tone burst or the corresponding frequency component of the complex stimulus. Corresponding results from subtractions of the two spectra (composite – complex) are depicted in the rightmost panels of Figs. 1 and 2. The portions that fall above or below the zero line indicate the amount by which the composite response was greater or less than the complex response. When there was no difference between the composite and complex responses, or when the response levels were

Fig. 2. Individual example, as in Fig. 1, but for an ear with relatively less correspondence between the two types of responses than for spectra in Fig. 1. Despite the poorer correspondence seen at several frequency regions, the general patterns of the responses are very similar. For both the complex stimulus and for the single tone burst at 3 kHz, the ear failed to respond in this frequency region with a stimulus level of 39 dB SPL.
below the noise floor, the datapoints then fall on the line crossing the ordinate at zero. The differences between the responses were relatively small, but were greatest at the higher level of stimulation. Although these two examples represent ears with (Fig. 1) and without (Fig. 2) SOAEs, the tendency for close correspondence of the two spectra was not biased uniformly towards ears with SOAEs. There were essentially no differences in the patterns of superposition between ears with and without SOAEs. Otherwise, the general trends illustrated by the results from these two ears were present in the mean results for the 28 ears, as depicted in Fig. 3.

The leftmost panels of Fig. 3 represent the mean spectra of the composite (solid lines) and complex (dotted lines) responses by stimulus level. These spectra bear a strong resemblance to each other both in configuration and amplitude. The mean of the differences between the composite and complex responses (composite - complex) are indicated by the solid lines in the rightmost panels of the figure. The dashed lines represent ±1 SD. As in the individual examples, positive and negative excursions of the line represent the relation of the composite to complex response. Differences in the amplitudes between the composite and complex responses were small. The range of the absolute values of the mean differences was from 0 to 3.4 dB for individual frequencies over all levels. There was no correlation between the magnitude of the differences and the absolute amplitudes of the responses. Although all responses from ears with SOAEs were generally greater in amplitude and contained more peaks and valleys, the superimposed composite and complex results were essentially indistinguishable from those without SOAEs. As illustrated both by the mean results in Fig. 3 and the individual results in Figs. 1 and 2, the primary differences between the two spectra were in the frequency regions associated with the high-frequency slopes of the 1- and 2-kHz peaks. These differences increased with stimulus level. However, direct comparisons between the level differences for the higher level stimulus ('nonlinear') and those obtained with the two lower levels (linear) are not possible because of differences in test methods. Overall, the composite spectra were higher in amplitude than were the complex spectra, especially for stimuli at 75 dB SPL. These trends may be confirmed by comparing the means from the subtractions of the two spectra (com-

![Fig. 3. Mean results from the complex and composite spectra for the 28 ears (leftmost panels) and the resulting mean differences from the subtraction of the two spectra (Composite - Complex). The solid lines on the rightmost panels represent the mean differences and the dashed lines are one standard deviation for the corresponding means. The main differences between the two spectra were in the frequency regions of the high-frequency slopes of the 1 and 2-kHz peaks for the higher level stimuli due to greater amplitudes of the components in the composite result. The differences in the spectra for the lowest level of stimulation are probably due to normal variability rather than strict departures from the superposition principles observed for the two higher stimulus levels.](image-url)
4. Discussion

Our results confirmed the general frequency-specific nature of tone burst evoked OAEs and further characterized the property of linear superposition. Emitted responses evoked by frequency-restricted transient stimuli are known to have energy peaks that correspond to the frequency components of the stimulus (e.g., Elberling et al., 1985; Harris and Probst, 1990; Norton and Neely, 1987; Probst et al., 1986; Kemp et al., 1990; Stover and Norton, 1991). This was clearly evident in the results of our study. The property of linear superposition of TEOAEs was tested specifically by the design of the stimuli that we used. Three tone bursts were first presented separately. The tone bursts were added digitally and presented as a complex stimulus. The TEOAEs resulting from the complex stimulus were compared to a digital addition of the separate TEOAEs produced by the individual tone bursts. Any differences between the two results were interpreted as departures from linear superposition.

The comparison of the averaged responses demonstrated nearly perfect linear superposition for the results obtained with a stimulus level of 39 dB (Fig. 3). Only slight deviations from linear superposition were present in the responses to the 57 dB stimuli and the deviations became more obvious when stimuli were at 75 dB. Therefore, linear superposition is maintained for relatively low levels of short-duration stimulation, such as those used in this study. More nonlinear interactions arise as stimulus level increases. The principal departure from linear superposition that we observed was a small reduction in the amplitude of the high-frequency portions of the two lower-frequency peaks produced by the complex stimulus (Fig. 3). The maximal averaged reduction was 3.4 dB. It has been shown in many physiological and psychophysical experiments (e.g., Abbas and Sachs, 1976; Houtgast, 1972; Shannon, 1976) that suppressors at frequencies within both lower and higher frequency boundaries of regions surrounding a target stimulus can reduce the amplitude of the target. For the TEOAEs that we measured, it appears that there is a small suppression of the lower frequency components produced by the higher frequencies in the stimulus. Similar findings for OAE suppression have been discussed by other investigators.

Brass and Kemp (1993) demonstrated that for stimulus-frequency otoacoustic emissions (SFOAEs), higher frequency stimuli suppress lower-frequency responses more easily than vice versa when small amounts of suppression and low level stimuli are involved. Zwicker and Wessel (1990) measured the rate of suppression for a TEOAE and obtained qualitatively similar results as did Brass and Kemp for SFOAEs. Higher-frequency suppressors started to suppress the response at lower levels than did lower-frequency suppressors, but the
rate of the suppression was more gradual than that from lower-frequency suppressors. These differences in suppression rate that were dependent upon the frequency of the suppressor were also reported by Kemp and Chum (1980) for SFOAEs.

Brass and Kemp (1993) proposed a model of the nonlinear interactions of the travelling waves associated with SFOAE suppression that fits well with our findings. The spectral changes that were observed during complex stimulation are also consistent with models proposed by Geisler, Yates, Patuzzi and Johnstone (1990) to account for two-tone suppression. In these models, a high-frequency stimulus (suppressor) presented at a level equal to or lower than a lower-frequency probe stimulus will force the generators in the basal regions into saturation. These generators will then no longer contribute to the energy in the response for the lower frequency probe. Figures in Geisler et al. (Fig. 11, 1990) and in Brass and Kemp (Fig. 1, 1993) illustrate this effect clearly. The primary 'active' areas for response peaks are located basal to the place of maximum displacement of the basilar membrane. The same nonlinear interactions that seem to be a property for continuous tonal stimulation may also occur for complex short duration stimuli.

The specific origin of OAEs is still not known. However, a general hypothesis is that there are generators (as yet undefined) distributed along the cochlear partition that are stimulated by a broadband transient signal and contribute energy differentially to the final result that is measured by a probe sealed in the outer ear canal (Avan et al., 1991; Norton and Neely, 1985). Although the majority of a TEOAE response to a tone burst is thought to be generated in the cochlear region tuned to the tone burst frequency, there is likely some contribution from generators located basally in the cochlea to that site (e.g., Avan et al., 1991; Brass and Kemp, 1993; Sutton, 1985). Therefore, with simultaneous presentation of a multi-frequency stimulus, the higher-frequency, or basal, components would be expected to saturate, which would result in small reductions of these contributions to the overall response pattern, such as observed in our results.

In summary, separate and combined tone burst stimuli are superposed linearly within the cochlea at low levels. A small amount of suppression of the high-frequency portions of the lower frequency tone bursts is present during multi-tone-burst stimulation. These findings fit well with current models of cochlear function. In addition, the technique that we have described allows noninvasive measurement of cochlear interactions for short duration stimuli that may be difficult to determine with other methods. The technique and its results may be useful in further testing predictions and performance of cochlear models.

5. Acknowledgements

We would like to thank David Kemp for his generous help in modifying the software of the ILO88 to perform the stimulus generation and analyses that we required. The comments of two anonymous reviewers were especially helpful in revising the manuscript.

Grant Support: Swiss National Foundation Projects Nr. 32-25514.88 and Nr. 32-32348.91

6. References


