Linear combination of auditory steady-state responses evoked by co-modulated tones

Guérit, François; Marozeau, Jeremy; Epp, Bastian

Published in:
Journal of the Acoustical Society of America

Link to article, DOI:
10.1121/1.5007757

Publication date:
2017

Document Version
Publisher's PDF, also known as Version of record

Linear combination of auditory steady-state responses evoked by co-modulated tones

François Guérit,[a] Jeremy Marozeau, and Bastian Epp
Hearing Systems group, Department of Electrical Engineering, Technical University of Denmark, DK-2800, Kongens Lyngby, Denmark
fguerit@elektro.dtu.dk, jemaroz@elektro.dtu.dk, bepp@elektro.dtu.dk

Abstract: Up to medium intensities and in the 80–100-Hz region, the auditory steady-state response (ASSR) to a multi-tone carrier is commonly considered to be a linear sum of the dipoles from each tone specific ASSR generator. Here, this hypothesis was investigated when a *unique* modulation frequency is used for all carrier components. Listeners were presented with a co-modulated dual-frequency carrier (1 and 4 kHz), from which the modulator starting phase Φ_i of the 1-kHz component was systematically varied. The results support the hypothesis of a linear superposition of the dipoles originating from different frequency specific ASSR generators.

[QJF]
Date Received: July 17, 2017 **Date Accepted:** October 2, 2017

1. Introduction

The auditory steady-state response (ASSR) is an auditory evoked potential which follows the repetition rate, defined by the modulation frequency f_m, of an ongoing sound signal (Picton *et al.*, 2003). For repetition rates between 80 and 100 Hz, the ASSR has been shown to arise from brainstem sources, while at lower rates (below 40 Hz), mostly sub-cortical and cortical sources are involved (Herdman *et al.*, 2002a).

When evoking the ASSR with sinusoidally amplitude-modulated (SAM) tones, an activation of auditory nerve fibers within a narrow region of the basilar membrane (Picton *et al.*, 2003) is assumed. The response to multiple SAM tone carriers with differing modulation frequencies has been shown to be a linear combination of the responses to each SAM component in the 80–100 Hz range of repetition rates (e.g., Herdman *et al.*, 2002b). However, for modulation frequencies around and below 40 Hz, multiple ASSR components do not combine linearly (John *et al.*, 1998), presumably because of interactions within the sub-cortical and cortical sources of the ASSR.

At higher stimulation levels, this linear combination of the ASSR components does not hold (Picton *et al.*, 2007). This can be explained by the nonlinear mechanics of the auditory periphery: a travelling wave excited by a pure tone carrier does not only result in an isolated vibration around the peak region of the carrier, but also evokes vibrations basal to that region. Stimuli presented at higher levels and composed of multiple frequency components are thus likely interacting across different regions along the basilar membrane. The contribution of different tonotopic regions to the ASSR has also been addressed in the context of chirp-evoked ASSRs (Elberling *et al.*, 2007), where it was found that the amplitude of the ASSR can be increased by stimulation with chirps accounting for the dispersion properties of the basilar membrane. For these stimuli, it is, however, not clear how each tonotopic region contributes to the measured ASSR other than that the overall amplitude increases.

For binaural stimulation with modulation frequencies around 80 Hz, a linear combination of ASSRs has been shown, also for components having the same modulation frequency, suggesting either the independence of two separate sources, or the linearity of a unique source of ASSR (e.g., Poelmans *et al.*, 2012).

Here, the assumption of a linear, monaural superposition of multiple co-modulated sources of ASSR in the 80 Hz region was investigated. The ASSR was recorded with electroencephalography (EEG), and was evoked by two SAM tones centred, respectively, at 1 and 4 kHz. Both carriers were modulated with the same modulation frequency but with a relative phase that was varied across conditions. It is

[a]Author to whom correspondence should be addressed.

hypothesized that the overall response measured using EEG is the vector sum of the ASSR evoked by each SAM tone separately, and will be sensitive to the relative modulator phase between the SAM tones.

The results will contribute to the understanding of how multiple sources of ASSR combine into the electrical signal measured at the scalp.

2. Methods

2.1 Subjects

Nine subjects participated in the experiment. Their hearing thresholds were below 20 dB hearing level at all audiometric frequencies (125 Hz to 8 kHz), and the mean age was 29.8 years, ranging from 25 to 40 years. The experimental procedure was approved by the Danish Science-Ethics Committee (ref. number H-3-2013-004), and written informed consent was obtained from all subjects before data collection.

2.2 Stimulus and apparatus

Seven different stimuli were used to elicit ASSRs, consisting of two SAM tones, s_{1k} and s_{4k}, and of five combined versions of those same tones with varied modulator starting phases of s_1k. The carrier frequencies of the two tones, f_{1k} and f_{4k}, were set, respectively, at 1 and 4 kHz. The carriers were 100% modulated at a frequency f_m of 88 Hz, as shown in Eqs. (1) and (2),

$$s_{1k}(t) = a_{1k}\sin(2\pi f_{1k}t)\left(\frac{1+\sin(2\pi f_m t+\Phi_i)}{2}\right), \tag{1}$$

$$s_{4k}(t) = a_{4k}\sin(2\pi f_{4k}t)\left(\frac{1+\sin(2\pi f_m t)}{2}\right). \tag{2}$$

When s_{1k} was presented in isolation, its modulator starting phase Φ_i was set to 0. For the five co-modulated conditions, stimuli were created by setting Φ_i to values distributed around the unit circle ($c^{\Phi_i} = s_{1k}^{\Phi_i} + s_{4k}$; $\Phi_i = 2i\pi/5$, $i = 0, 1, \ldots, 4$), while s_{4k} was kept the same.

To avoid distortions in the co-modulated conditions, the two carriers were played separately through two ER-2 earphones mounted on an ER-10B+ probe (Etymotic Research, Inc.), and connected to the computer through a Phonitor mini amplifier (SPL electronics GmbH) and a Fireface UCX sound card (Audio AG). Both a_{1k} and a_{4k} were adjusted to deliver s_{1k} and s_{4k} at 65 dB sound pressure level (SPL) in isolation, using a B&K 4137 ear coupler and a B&K 2636 sound level meter (Brüel & Kjær A/S).

2.3 ASSR recording and analysis

Subjects were seated in a double-walled, electrically shielded, sound-attenuating booth. They were instructed to relax and stay calm. They watched a silent film with subtitles throughout the whole recording session, and were awake at all time. The stimulated ear was randomized across subjects, and the opposite ear was occluded with an ear plug, to avoid acoustical cross-talk.

EEG signals were recorded using a BioSemi ActiveTwo system (Biosemi B.V.), sampled at 8192 Hz, and analyzed offline with MATLAB (The MathWorks, Inc.).

A vertical electrode montage was used, using the 10/20 system, with two electrodes: P9 or P10 at the left or right mastoid, respectively, and Cz at the vertex. If the right ear was stimulated, the difference between Cz and P10 was computed, while Cz and P9 were used for the left ear stimulation. Each stimulus condition was recorded for approximately 10 min (608 s). The signal was cut into epochs of 16 s, and any epoch exceeding 80 μV was discarded from the processing. A weighted averaging method based on the standard deviation in each epoch (John *et al.*, 2001) was then applied to obtain a single 16-s epoch, from which the fast Fourier transform (FFT) was computed with a bin width of 0.0625 Hz. A F-ratio was computed between the power of the FFT bin at 88 Hz (chi-squared variable with 2 degrees of freedom) and the power of the EEG background noise (96 neighbouring bins, ± 3 Hz, 96×2 degrees of freedom). The ASSR was deemed above the noise floor when the null hypothesis that both noise and ASSR component came from the same F distribution was rejected ($p \leq 0.01$, Dobie and Wilson, 1996). This corresponds to a signal-to-noise ratio above or equal to 6.73 dB [$= 10\log_{10}(P_{signal+noise}/P_{noise})$].

Due to anatomical differences (head size, neural sources), inter-subject variability in the group delay (and therefore the phase) is expected. Because of this, co-modulated responses are likely to be in/out of phase for different values of Φ_i across listeners. Measured amplitude responses to the c^{Φ_i} stimuli were therefore shifted to have their maximum value at $\Phi_i = 0$ rad (Riecke *et al.*, 2015).

Before computing the phase of the co-modulated ASSRs, the response to s_{4k} in isolation was subtracted ($\mathbf{ASSR}(c^{\Phi_i}) - \mathbf{ASSR}(s_{4k})$). In the case of linearity, the phase of this vector subtraction should therefore equal Φ_i, the phase of s_{1k}. Again, to account for inter-subject variability in group delay, the phase of the afore-mentioned subtraction was shifted to be 0 rad for Φ_0. Unless specified, the Φ_0 condition was removed from all statistical analysis, as data for this point do not satisfy independence requirements.

3. Results

3.1 Responses to single carriers

The amplitudes of the ASSRs were above the noise floor for 8 out of 9 subjects in response to s_{1k}, and 7 out of 9 in response to s_{4k}. They were similar in amplitude [Fig. 1(B)] and comparable in value to previously reported amplitudes at those stimulation levels (Picton *et al.*, 2007). Overall (including the responses to co-modulated carriers), mean amplitude and standard deviation of the significant ASSRs were, respectively, 52.5 and 26.5 nV. One subject had higher noise levels (mean/s.d. of 48.9/4.3 nV versus a mean and s.d. of, respectively, 10.7 and 2.5 nV for the other subjects). Since this subject did show significant ASSRs in some conditions, it was not excluded. We, however, controlled for every statistical analysis that removing this subject did not change the main conclusions.

The difference in phase between the responses to s_{1k} and s_{4k} was 103°, as shown in Fig. 1(A). In order to link the ASSR phase to an estimate of cochlear travel time, and assuming a linear phase of the frequency components along the cochlea, this phase corresponds to a latency difference of 3.3 ms for a modulation frequency of 88 Hz ($103/360 \times 88$). This difference was statistically significant (paired-sample t-test with the seven subjects having both responses above significance; df = 6; p = 0.0019, t = 5.2691, 95% confidence interval = 1.7–4.8 ms).

3.2 Co-modulated responses

By combining Eqs. (1) and (2), one can hypothesize that the mean vector sum of all co-modulated responses should equal the response to s_{4k} in isolation, as shown in Eqs. (3)–(5) and Fig. 1,

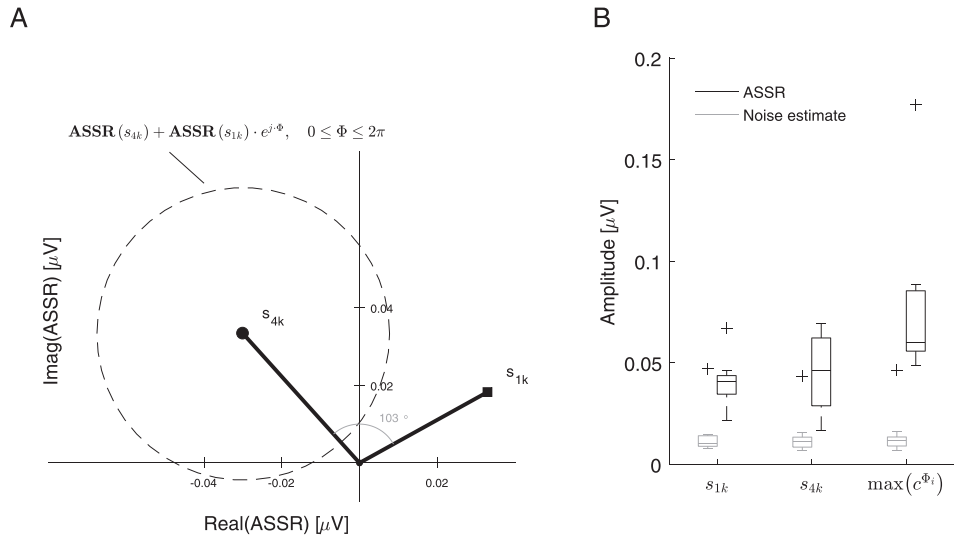

Fig. 1. (A) Averaged complex-valued ASSRs across subjects in response to the s_{1k} (square) and s_{4k} (circle) stimuli. Shown with a dashed line is the expected co-modulated response (linear vector summation) when varying Φ_i, the modulator starting phase of s_{1k}. (B) Across-subject distribution of ASSR amplitudes (black) and of noise estimates (grey) in response to the s_{1k}, s_{4k}, and c^{Φ_i} stimuli. For c^{Φ_i}, only the maximum amplitude across all angles is shown. *Lower and Upper limits of the boxes:* 25th and 75th percentiles. *Horizontal line:* median. *Whiskers:* 25th (or 75th) percentile minus (or plus) 1.5 the interquartile range.

$$\frac{1}{5}\sum_{i=0}^{4}\mathbf{ASSR}(c^{\Phi_i}) = \frac{1}{5}\sum_{i=0}^{4}\left(\mathbf{ASSR}(s_{4k}) + \mathbf{ASSR}(s_{1k})e^{j(2i\pi/5)}\right) \quad (3)$$

$$= \mathbf{ASSR}(s_{4k}) + \mathbf{ASSR}(s_{1k})\sum_{i=0}^{4} e^{j(2i\pi/5)} \quad (4)$$

$$= \mathbf{ASSR}(s_{4k}). \quad (5)$$

A t-test comparing the mean vector sum of all co-modulated responses to the response to s_{4k} in isolation showed no significant difference (df = 6, $p_{real} = 0.7617$, $t_{real} = -0.3173$, $p_{imaginary} = 0.5832$, $t_{imaginary} = -0.5798$, Pearson's $r = 0.8322$ when pooling real and imaginary values, $r = 0.8310$ when removing the subject with high noise level). This t-test excluded two subjects who had missing data in one condition (hence df = 6), and was run on both real and imaginary parts of the ASSR, as they can be considered to be independent variables (Dobie and Wilson, 1996).

Amplitude. As shown in Fig. 1(B), the individual maximum for each subject across all co-modulated conditions was significantly larger than the response to the single carriers in isolation (pairwise t-Tests, paired data within subjects, Bonferroni corrections, p = 0.00023 and p = 0.000232 for s_{1k} and s_{4k}, respectively). The data were log-transformed for this test to account for the presence of a subject with higher overall amplitudes. Figure 2(A) shows the amplitudes obtained for all co-modulated stimuli, with the individual responses aligned to be largest at $\Phi_0 = 0$ rad, and normalized by their value at Φ_0. A multilevel approach for repeated measures was employed, with the subjects as a random factor (Field *et al.*, 2012), and failed to show a significant effect of Φ_i on the relative amplitude as plotted in Fig. 2(A) [$\chi^2(3) = 6.98$, p = 0.0726, data points below the noise floor excluded]. By adding the points below the noise floor in this statistical analysis, the effect of Φ_i becomes significant [$\chi^2(3) = 13.67$, p = 0.0034]. It is worth noticing that conditions Φ_2 and Φ_3 had the highest number of recordings below the noise floor [Fig. 2(B)], where s_{1k} and s_{4k} were expected to be out of phase.

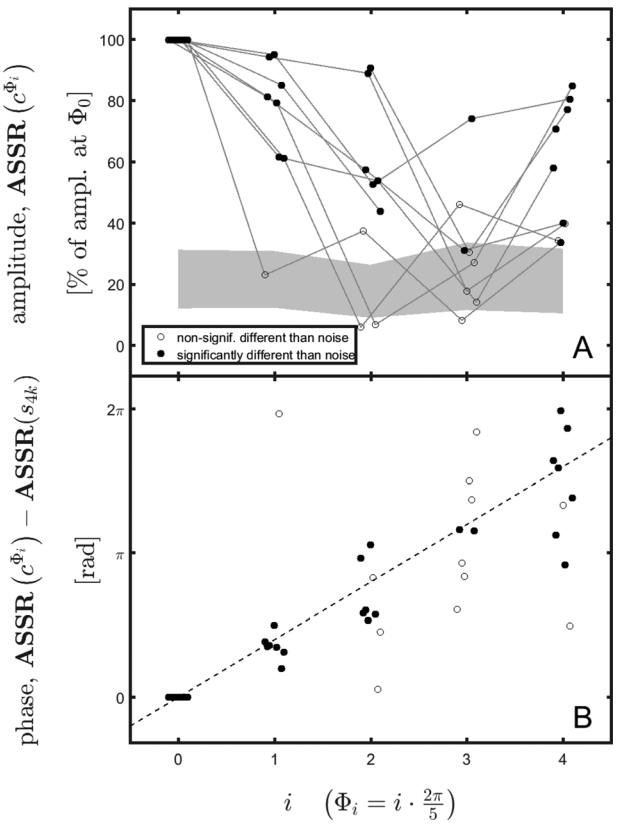

Fig. 2. (A) Amplitudes of individual co-modulated responses, shifted to have their maximum at Φ_0, and normalized by their value at Φ_0. A small jitter has been added to the x axis to improve readability, and the noise floor excursion (min to max) is shown with the grey shaded area. As noise levels differed across subjects, it can be seen that the amplitude (on a relative scale) required to have a significant response varies across subjects. (B) Phase of the vector subtraction $\mathbf{ASSR}(c^{\Phi_i}) - \mathbf{ASSR}(s_{4k})$ (which should be Φ_i in case of linearity), normalized to be 0 at Φ_0 and wrapped between 0 and 2π.

To assess whether the proportion of significant points was the same between different Φ_i conditions, a Cochran Q test was used, and showed a significant effect of Φ_i ($Q = 8.4000$, $df = 3$, $p = 0.0384$).

Phase. Figure 2(B) shows the phase values corresponding to Fig. 2(A), re-referenced to be 0 for Φ_0 and wrapped between 0 and 2π. Φ_i had a significant effect on the ASSR phase [$\chi^2(3) = 37.8$, $p < 0.0001$]. This effect was well fitted by a linear regression (intercept = –0.1255, slope = 0.9967, 95% confidence interval = 0.75–1.19, $r^2 = 0.7956$).

4. Discussion

In the 80–100-Hz range, it has been hypothesized that the ASSR evoked by SAM tones with different modulation frequencies is the linear superposition of the response to the SAM tones presented alone. This assumption has been shown to hold true if the carrier frequencies are separated at least by an octave, and if medium levels are used (Herdman *et al.*, 2002b). The present study supports the hypothesis of a linear superposition, and expands it to the case of carriers modulated with a unique modulation frequency presented monaurally (this has already been shown binaurally, e.g., in Poelmans *et al.*, 2012). However, because the ASSR measured by EEG is a gross potential, it cannot be distinguished whether the observed effects in the presented paradigm are due to a superposition of two independent sources contributing to the ASSR or if the effects are due to neural interactions within a single source of the ASSR.

4.1 Linearity of the co-modulated conditions

Under our linearity assumption, and because Φ_i was evenly distributed around the unit circle, summing all co-modulated responses should not be significantly different than the response to s_{4k} in isolation, and this is indeed what we could see in our recordings.

Additionally, manipulating the modulator starting phase of s_{1k} in the co-modulated conditions had a significant effect on both phase and amplitude of the ASSR. This effect was consistent with a linear sum when analyzing the co-modulated phase response [Fig. 2(B)], while the individual patterns of the amplitude were more variable [Fig. 2(A)]. These deviations seen in the amplitude of the co-modulated responses might be due to the inherent test-retest variability of the ASSR. Finally, when both single carriers were supposedly out of phase (conditions Φ_2 and Φ_3 in Fig. 2), it was often impossible to record a significant response, even with 10 min of recording and median noise levels of 10.9 nV.

Taken together, these results support the hypothesis of a linear superposition, and that distinct neural populations are represented in the ASSR, even when using a unique modulation frequency. John *et al.* (2003) measured the ASSR of 4 SAM tones at 0.5, 1, 2, and 4 kHz, co-modulated and in isolation. The co-modulated response was 25% lower than expected by a linear vector superposition of the responses in isolation. This reduction, not seen in our study, might be explained by the fact that they used four carriers separated by only one octave (while we used only two carriers separated by two octaves). This might have led to interactions at the level of the basilar membrane, such as mutual suppression.

In contrast to multi-tone carrier ASSRs, single-evoked ASSR growth functions do not show a saturation for stimulus levels above 60 dB SPL (Picton *et al.*, 2007). Based on the results of the present study, one might however speculate that responses evoked by off-frequency regions also contribute to the measured amplitude in single-evoked ASSRs, and that the measured ASSR is a linear combination of responses evoked by on- and multiple off-frequency regions with different relative phase.

4.2 Further use of this paradigm

As linearity seems to be respected with this paradigm, any measured non-linearity could be used as a marker for envelope interactions at the level of the cochlea. An example is for cochlear implant users, where the spread of electrical current produces marked envelope interactions in a behavioural task (Galvin *et al.*, 2015).

5. Conclusions

This study suggests that the ASSR at 88 Hz with co-modulated carriers presented monaurally is a linear sum of the response to each carrier, as supported by the phase behaviour of the co-modulated response and the vector sum of all co-modulated responses.

Such a paradigm, where the phase difference between co-modulated carriers is varied, is therefore suitable for analyzing envelope interactions with a unique modulation frequency and at peripheral levels of the auditory system.

Raw data files (.bdf and stimuli files) are accessible on Zenodo (Guérit *et al.*, 2017).

Acknowledgments

F.G. is supported by the Oticon Centre of Excellence for Hearing and Speech Sciences (CHeSS) and the Technical University of Denmark. The authors would like to thank Gerard Encina Llamas for the ASSR analysis code and James Harte for fruitful discussions. The authors would also like to thank two anonymous reviewers for constructive comments on a previous version of this manuscript.

References and links

Dobie, R. A., and Wilson, M. J. (**1996**). "A comparison of *t* test, *F* test, and coherence methods of detecting steady-state auditory-evoked potentials, distortion-product otoacoustic emissions, or other sinusoids," J. Acoust. Soc. Am. **100**, 2236–2246.

Elberling, C., Don, M., Cebulla, M., and Stürzebecher, E. (**2007**). "Auditory steady-state responses to chirp stimuli based on cochlear traveling wave delay," J. Acoust. Soc. Am. **122**, 2772–2785.

Field, A., Miles, J., and Field, Z. (**2012**). *Discovering Statistics using R* (Sage, Thousand Oaks, CA).

Galvin, J. J., Oba, S. I., Başkent, D., Chatterjee, M., and Fu, Q.-J. (**2015**). "Envelope interactions in multichannel amplitude modulation frequency discrimination by cochlear implant users," PLoS One **10**, e0139546.

Guérit, F., Marozeau, J., and Epp, B. (**2017**). "Supplementary data: Co-modulated auditory steady-state responses," Zenodo.

Herdman, A. T., Lins, O., Van Roon, P., Stapells, D. R., Scherg, M., and Picton, T. W. (**2002a**). "Intracerebral sources of human auditory steady-state responses," Brain Topogr. **15**, 69–86.

Herdman, A. T., Picton, T. W., and Stapells, D. R. (**2002b**). "Place specificity of multiple auditory steady-state responses," J. Acoust. Soc. Am. **112**, 1569–1582.

John, M. S., Dimitrijevic, A., and Picton, T. W. (**2001**). "Weighted averaging of steady-state responses," Clin. Neurophysiol. **112**, 555–562.

John, M. S., Dimitrijevic, A., and Picton, T. W. (**2003**). "Efficient stimuli for evoking auditory steady-state responses," Ear Hear. **24**, 406–423.

John, M. S., Lins, O. G., Boucher, B. L., and Picton, T. W. (**1998**). "Multiple auditory steady-state responses (MASTER): Stimulus and recording parameters," Audiology **37**, 59–82.

Picton, T. W., van Roon, P., and John, M. S. (**2007**). "Human auditory steady-state responses during sweeps of intensity," Ear Hear. **28**, 542–557.

Picton, T. W., John, M. S., Dimitrijevic, A., and Purcell, D. (**2003**). "Human auditory steady-state responses," Int. J. Audiol. **42**, 177–219.

Poelmans, H., Luts, H., Vandermosten, M., Ghesquiere, P., and Wouters, J. (**2012**). "Hemispheric asymmetry of auditory steady-state responses to monaural and diotic stimulation," J. Assoc. Res. Otolaryngol. **13**, 867–876.

Riecke, L., Formisano, E., Herrmann, C. S., and Sack, A. T. (**2015**). "4-Hz transcranial alternating current stimulation phase modulates hearing," Brain Stimul. **8**, 777–783.