ELECTROPHYSIOLOGIC CORRELATES OF SPATIAL RELEASE FROM MASKING

By

KAILYN ALYSSA MCFARLANE

A Thesis Submitted to The Honors College

In Partial Fulfillment of the Bachelors degree With Honors in

Neuroscience and Cognitive Science

THE UNIVERSITY OF ARIZONA

M A Y 2 0 1 8

Approved by:

Dr. Barbara Cone, Ph.D.

Department of Speech, Language, and Hearing Sciences

ABSTRACT

Speech is difficult to understand in the presence of background noise. When the speech signal and source(s) of the background noise are spatially separated, it becomes easier to detect the speech. This is known as spatial release from masking (SRM). Previous research using perceptual test methods has demonstrated that listeners with hearing loss have variable benefit from SRM. This study documented the benefits of SRM using the cortical auditory evoked potential (CAEP) in response to speech tokens as noise level and location were varied. CAEPs from twenty normally hearing adults were recorded in response to consonant-vowel speech tokens in quiet, co-located noise and spatially separated noise. SRM benefit was measured by comparing the latency an amplitude of CAEP components P1, N1, and P2 in co-located and spatially separated conditions. Psychophysical tests of speech perception were completed in the same co-located and spatially separated noise conditions. Latencies and amplitudes of the CAEP components showed a systematic shift as a function of noise location, level, SNR, and stimulus. Co-located conditions across all stimuli, levels, and SNRs had longer latencies and smaller amplitudes than the spatially separated conditions, demonstrating an electrophysiologic analog of perceptual SRM. These results provide a baseline for investigation of SRM benefits in adults with hearing loss.

INTRODUCTION

It is a common human experience to have difficulty discerning speech in noisy environments, such as a crowded restaurant. Although our brains do an admirable job in auditory scene analysis (Bregman, 1990) and parsing out important signals from irrelevant noise, the noise can still prevent the signal from being heard. This is referred to as masking. When a speech signal and noise source are co-located (at the same location on the horizontal, or azimuth, plane), the speech may be completely masked by the noise. However, this masking effect can be diminished if the speech and noise are spatially separated (Gelfand, 2009). This is termed spatial release from masking (SRM). It has been well established that the spatial separation of speech and the competing noise lead to an improved ability to detect and discriminate speech (Bronkhorst, 2000).

The benefit provided by SRM is affected by multiple factors, including hearing ability (Marrone, Mason, & Kidd, 2008). Previous research has demonstrated that individuals with hearing loss have great difficulty hearing speech in noise and do not experience benefits from SRM in comparison to individuals with normal hearing (Ching et al., 2011). This knowledge has been established using psychophysical (i.e. perceptual) methods (Marrone et al., 2008; Ching et al., 2011).

The aim of the present study was to document the benefits of spatial release from masking using electrophysiologic methods. This was done by recording the cortical auditory evoked potential (CAEP) in response to speech tokens as noise level and location were varied. The CAEPs produced by these different conditions allowed us to evaluate the underlying brain mechanisms of the difficulties listeners have in noise and how those difficulties can be remedied by the manipulation of noise characteristics (e.g. level and location in relation to the speech signal). Once these speech-in-noise processing mechanisms are known in individuals with normal hearing, the study can be extended to individuals with hearing loss and the CAEPs of the two populations can be compared. Knowledge of these mechanisms are important in developing strategies to help individuals with hearing loss navigate noisy environments. Examples of such strategies include designing new signal processing algorithms for hearing aids and developing listening therapies that may improve speech perception in noise abilities.

In this study, we specifically investigated the effect of (1) voiced vs. unvoiced consonant-vowel tokens, (2) level, (3) signal-to-noise-ratio (SNR) and (4) co-located vs. spatially separated noise on CAEP latency and amplitude. We also investigated the effect of SNR and noise location on speech perception in noise as measured using standardized tests. We hypothesized that the CAEP could be used as a bio-marker for brain mechanisms underlying speech-in-noise perception, because the CAEP represents the first stages of auditory processing at the cortical level. The predictions following from this hypothesis are:

1) CAEP latency differences would be observed between /da/ and /ta/ responses, because CAEP latencies are known to systematically shift with an increase in voice onset time (Sharma and Dorman, 1999; Almeqbel, 2016).
2) CAEP latency would decrease and amplitude would increase as stimulus level increased, because of the known stimulus intensity effects on the late auditory evoked potential (Hall 1991; Picton, 2010).
3) CAEP latency will decrease and amplitude will increase as SNR increases, because of the known SNR effects on CAEP (Billings et al., 2009; Martin & Stapells, 1999).

4) CAEP latency will decrease and amplitude will increase as the noise is moved from co-located to spatially separated condition, because of the known benefit of spatial release from masking (Bronkhorst, 2000; Gelfand, 2009).
5) Performance in speech-in-noise testing would improve with spatial separation of noise, because of the known improvement in speech intelligibility with spatial separation (Bronkhorst, 2000) and these differences will be correlated with CAEP latency and amplitude results.

METHODS

All methods were approved by the University of Arizona Human Subjects Protection Program (Institutional Review Board).

Participants

Twenty adults participated in the experiment, 2 of them male, with a mean age of 24 years (range 20-36 years). All participants met the inclusion criteria of bilateral hearing thresholds within normal limits (<20 dB HL) from 250-8000 Hz as determined by a pure tone hearing screening test.

Stimuli

The stimuli used to evoke the CAEP were synthesized consonant-vowel /da/ or /ta/. The overall duration of the tokens was 205ms. Shown in Table 1, the fundamental frequencies of the vowel portions of the tokens ranged from 98-140 Hz. The five vowel formants of the /ta/ stimulus ranged from 861-3445 Hz, and for /da/ the range was 732-3828 Hz. 10-talker noise babble, extracted from the AZBio recording, was used as the noise masker.

Speech tokens were presented at a rate of 1.5/s. Speech tokens were presented through loudspeakers at 0° azimuth and noise was presented through loudspeakers positioned at 0° and 90° azimuth in the co-located condition and spatially separated condition, respectively (see Figure 1). The token presentation levels were 15, 30, 45 and 60 dB SL re: psychophysical threshold for the speech token, and noise levels were at +10, 0 and -5 dB SNR. There were a total of 28 conditions, shown in appendix C.

Speech stimuli and noise maskers were calibrated using a Larsen Davis Model 824 sound pressure level meter with a half-inch microphone at the position of the head. An A-weighting was used for these measures.

Speech tokens were presented via the IHS Smart-EP system, amplified by a Crown D-75A amplifier routed to the JBL Control 1x model loudspeaker. Noise was presented via the Otometrics Madsen Astera audiometer to a NHT SuperOne model loudspeaker.

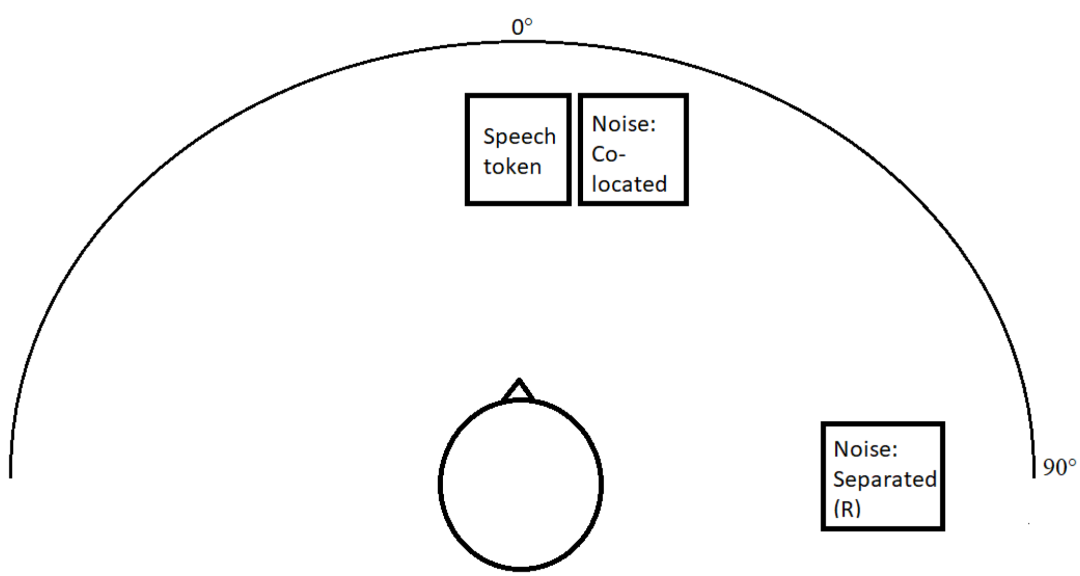


Figure 1. The loudspeaker positions for the co-located and spatially separated CAEP recording and speech-in-noise test conditions. The boxes represent the loudspeakers, labeled with their corresponding output and the arch represents the azimuth plane.

	/DA/	/TA/
Stimulus file duration	206.275 ms	205.193
Consonant duration	9 ms	47.461
Vowel duration	174 ms	157.732
Pitch (start - finish)	109.1 – 102.1 Hz	140.4-97.6 Hz
Formants (Hz):		
F1	732	861
F2	1335	1292
F3	2498	2326
F4	3058	3230
F5	3828	3445

Table 1. Spectral information of IHS /da/ and /ta/ speech stimuli.

Electrophysiologic Testing

CAEPs were obtained using Intelligent Hearing Systems (IHS) Smart-EP System. Electrodes were placed at Cz (vertex, non-inverting), A2 (right earlobe, inverting) and FpZ (forehead, ground). Each site was cleansed with a gentle exfoliant (Nu-Prep) before the electrodes were secured with paste and paper tape. Electrode impedances were kept under 10kΩ. The EEG was band-pass filtered at 1-30Hz and amplified with a gain of 94 dB. CAEPs were averaged over 200-300 sweeps, with a sampling rate of 1000Hz. The responses were averaged over a 500ms epoch following the onset of each token.

Behavioral Testing

The AzBio (see appendix A) and QuickSIN (see appendix B) tests were used to assess speech perception in noise. The AzBio target sentences were presented at 15 and 45 dB SL re: pure-tone average, at 10 and -5 dB SNR, in co-located and spatially separated 10-talker babble noise. All 8 AzBio lists, consisting of 20 sentences each, were used (e.g. List 1 presented at 15 dB SL, 10 SNR, co-located, List 2 presented at 15 dB SL, -5 SNR, co-located, etc.). Participants were instructed to listen to the sentences presented through the loudspeaker and repeat them back as verbatim as possible, and were encourage to guess when they were unsure. Responses were scored on the number of correct words in each sentence the participant correctly repeated. The percent correct (total words correct/total words possible) was calculated for each list. The QuickSIN sentences were presented at 15, 30, and 45 dB SL, in co-located and spatially separated babble noise. The QuickSIN test consists of six sentences of varying length with five target words each, with the SNR decreasing in 5 dB steps after each sentence, starting at an SNR of +25 dB, ending at 0 dB SNR). Participants were instructed to listen to the sentences presented through the loudspeaker and repeat them back as verbatim as possible, and to guess when unsure. Responses were scored on the basis of the number of target words in each sentence the participant correctly repeated. The total words correct was subtracted from 25.5 to determine the SNR at which the participant could repeat 50% of the words correctly. This calculation is termed "SNR loss" and was calculated for each list and noise condition.

Test Session

Participants completed all electrophysiologic and behavioral testing in two separate sessions of approximately two hours each. Sessions began and ended with electrophysiologic testing. Behavioral testing was administered halfway through each session to provide the participant a chance to reboot their attention and energy level, as participants would become sleepy and less engaged during the passive listening CAEP recordings. All testing took place in a sound treated audiometric test booth. For electrophysiologic testing, participants were tested while awake and seated in a chair. They were allowed to read, work on their computer, or look at their phone as long as their head maintained the correct orientation on the azimuth plane.

The first test session began with the hearing screening. The participant's threshold for the speech token was then determined using a stair-case procedure by decreasing in 10 dB steps and increasing in 5 dB steps until a 50% correct detection was obtained over 3 out of 4 test trials. The participant was then prepped for CAEP recording. The first CAEP recordings administered were speech in quiet at all four levels: 15, 30, 45, and 60 dB SL. After that, the level order was randomized for each participant (see Appendix C for a randomized condition table). After about 8 CAEP recording conditions, the participants performed the AzBio sentence repetition test. CAEP recording was resumed afterwards and continued until the session was over. The second session started with the third portion of the CAEP recording, then the QuickSIN test was administered, and then the remaining CAEP recording conditions were completed.

Data Analyses

CAEP latency and amplitude measures were made using rule-bound visual detection methods (Wunderlich and Cone-Wesson, 2001) by the lead examiners (KM & AP) who had been instructed by principal investigator (BC). The rules for visual detection of CAEP components included amplitude and latency criteria for each component, based upon values derived from the published literature on adult CAEPs (Picton, 2010). Based on the typical latency of a young adult for a moderately intense stimulus, the P1, N1, and P2 components are found at 50-80ms, 90-120ms, and 160-200ms, respectively. Amplitudes of 0.5μV or less were not considered to be responses. CAEP component peaks P1, N1, P2, and N2 were marked using the IHS software with a cursor to determine peak latency, and amplitudes were calculated as the difference between the peak and succeeding trough, or trough-to peak (e.g., P1-N1, N1-P2, P2-N2). Descriptive and inferential statistics were performed using Statview V.5.

RESULTS

Repeated measures analyses of variance were used to determine the effects of consonant voicing, stimulus level, SNR and noise location on CAEP component (P1-N1-P2) latencies and amplitudes. Each of these factors resulted in statistically significant effects on CAEP latency at the p<.001 level. SNR and noise location resulted in statistically significant effects on CAEP amplitude at the p<.001 level. The result of the analyses of variance are summarized in Table 2. The effects of SNR and noise location on speech perception in noise as measured using standardized tests AzBio and QuickSIN were also determined.

Effect of voiced vs. unvoiced consonant-vowel tokens on CAEP latency

Statistically significant differences in latencies were found between /da/ and /ta/ at each CAEP component, which are shown in Figure 2. Latency differences between /da/ and /ta/ were on the order of 8ms for P1, 16ms for N1, and 26ms for P2. The voiced consonant-vowel speech token /da/ evoked shorter CAEP P1-N1-P2 latencies than the unvoiced /ta/ token. Figure 3 shows the grand mean waveforms for /da/ and /ta/.

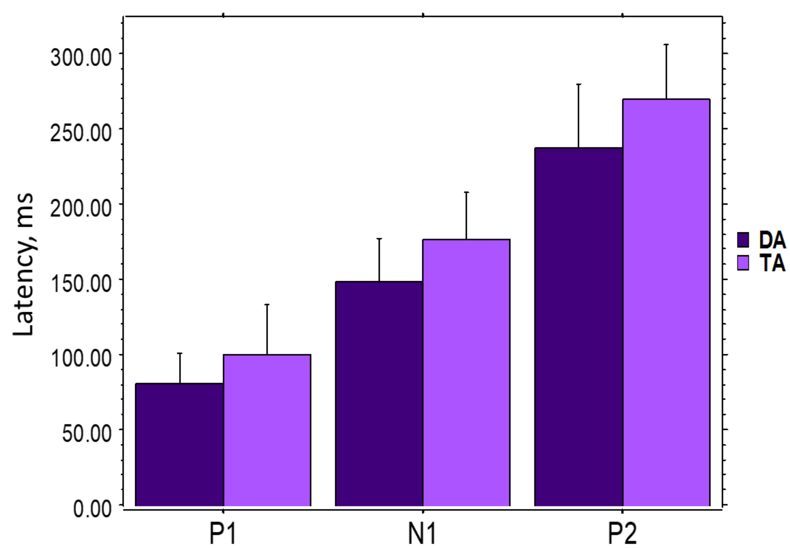


Figure 2. Latencies of CAEP components P1, N1, and P2 for /da/ and /ta/. Error bars indicate standard deviation.

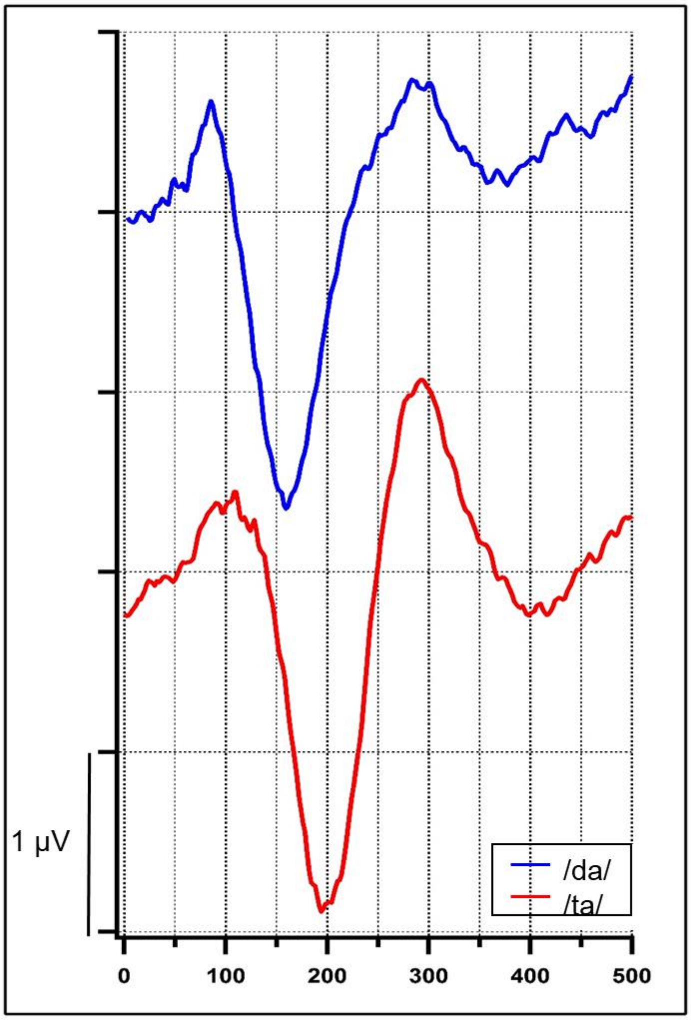


Figure 3. Grand mean averages of /da/ and /ta/ CAEPs. Data used for /da/ and /ta/ were averaged over all four stimulus levels and over 0 dB and -5 dB SNR, co-located.

Effect of level on CAEP latency and amplitude

CAEP latencies increased as levels were decreased (Figure 3). The greatest latency shifts were observed between the 15 dB SL and 30 dB SL conditions, and were on the order of 35ms for P1, 40ms for N1, and 40ms for P2.

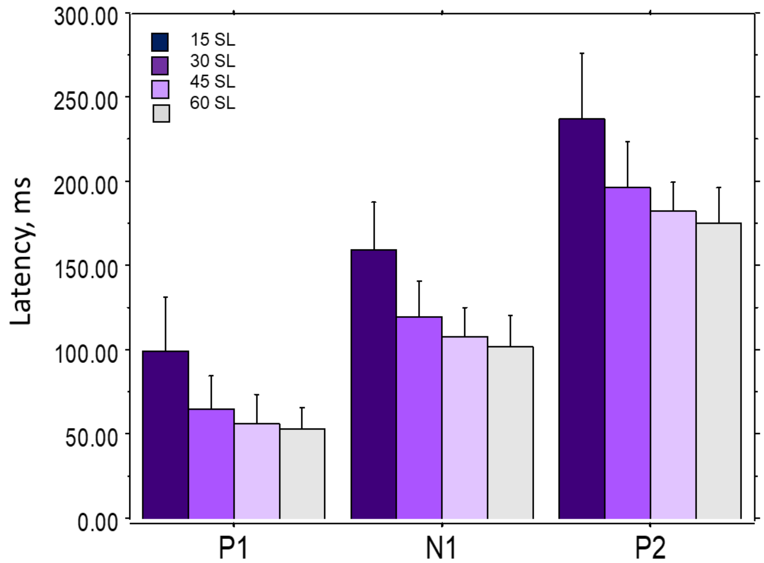


Figure 3. Average P1, N1, and P2 latencies at all four levels: 15, 30, 45, and 60 dB SL. As level increased, the latencies of each CAEP component decreased. The latencies of each CAEP component were averaged in the quiet conditions. Error bars indicate standard deviation.

CAEP amplitudes were measured as a function of level in quiet conditions. There was no significant change in amplitude as a function of level. As can be seen in Figure 4, amplitudes and morphology are stable at 30dB SL and above.

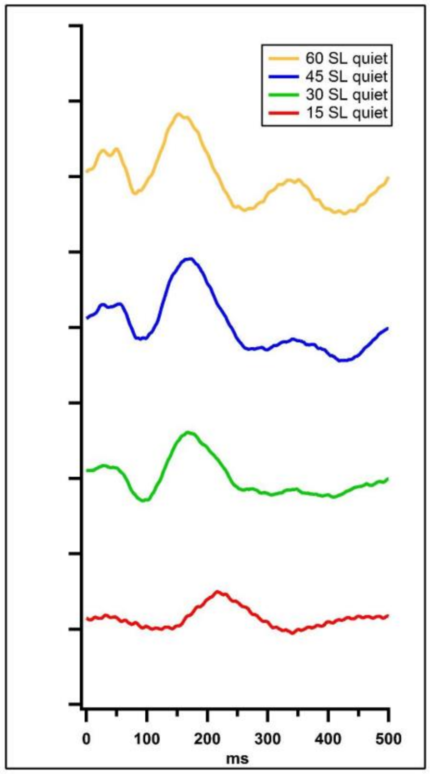


Figure 4. Grand mean averages of the CAEP waveforms for each level condition. These waveforms were recorded in response to the speech tokens at each level in the absence of noise (i.e. in quiet). The waveforms for /da/ and /ta/ at each level are averaged together in this figure.

As shown in Figure 4, the morphology of the waveform and the CAEP components became more distinguishable as the level increased. The P1, N1, and P2 components are all present at the 45 dB SL and 60 dB SL conditions, while the P1 component becomes less defined at 30 dB SL and only N1-P2 is evident at 15 dB SL.

Effect of signal-to-noise-ratio (SNR) on CAEP latency and amplitude

CAEP latencies decreased and CAEP amplitudes increased with increasing SNR (i.e. latencies were shorter and amplitudes were larger with improving SNR), shown in Figure 5 and Figure 6, respectively. The amount of latency shift observed from -5 dB to +10 dB SNR is larger than that seen for the latency shifts between 30 and 60 dB SL (in quiet).

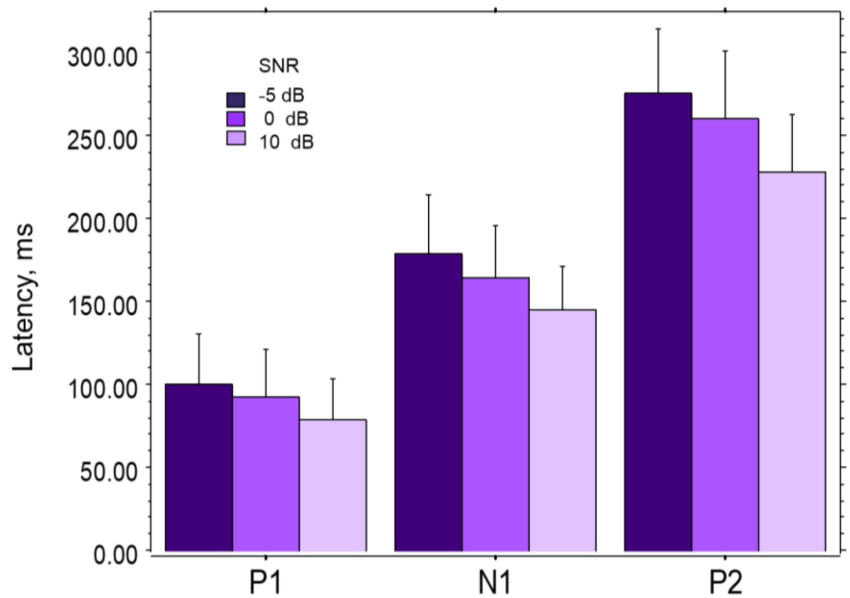


Figure 5. Average P1, N1, and P2 latencies for all three SNR conditions: -5, 0, and +10 dB. As SNR improved (increased from -5 to +10 dB), latencies for each CAEP component decreased. Latencies for each CAEP component at each SNR was averaged across all levels and noise location conditions. Error bars indicate standard deviation.

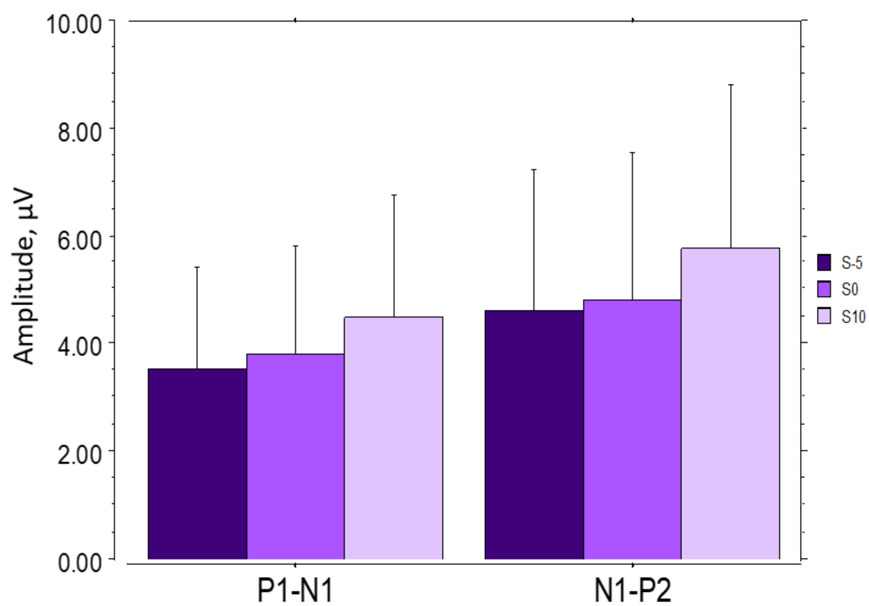


Figure 6. P1-N1 and N1-P2 amplitudes at all three SNR conditions. As SNR improved (increased from -5 to +10 dB), amplitudes of P1-N1 and N1-P2 increased. Amplitudes were averaged across all levels and noise location conditions. Error bars indicate standard deviation.

Effect of noise location on CAEP latency and amplitude

As the noise was moved from the co-located condition at 0° azimuth to the spatially separated condition at 90° azimuth, latency decreased (Figure 7) and amplitude increased across all CAEP components. The increase in N1-P2 amplitude due to spatial separation (Figure 9) is equivalent to the N1-P2 amplitude increase caused by increasing the SNR from -5 dB to +10 dB (Figure 6). Figure 8 illustrates the latency shift and amplitude changes that occur at each level as the noise moves from co-located to spatially separated.

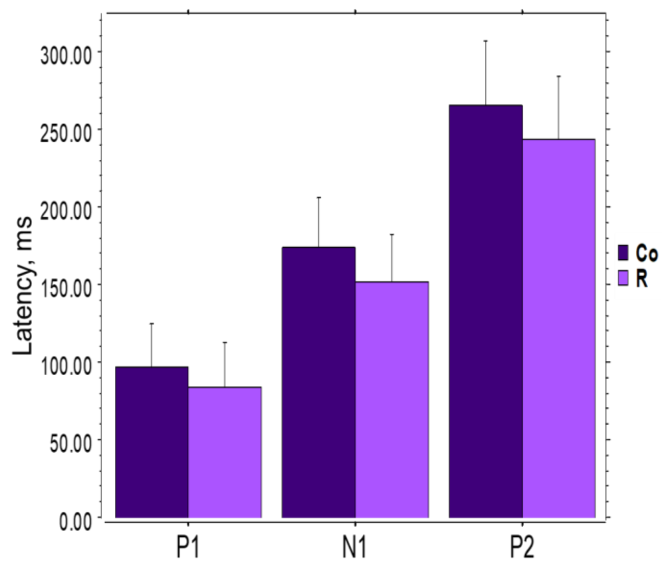


Figure 7. Averaged latencies of each CAEP component in the Co-located (Co) and spatially separated (R) conditions. The latencies for each CAEP component were averaged across all SNRs and all levels. Error bars indicate standard deviation.

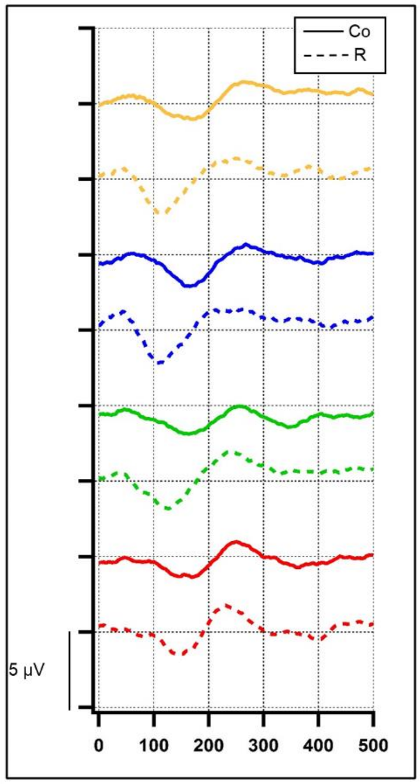


Figure 8. Grand mean averages of CAEP recordings of the co-located and spatially separated conditions at each level condition. Levels are organized from top to bottom in descending order (yellow represents 60 dB SL, red represents 15 dB SL).

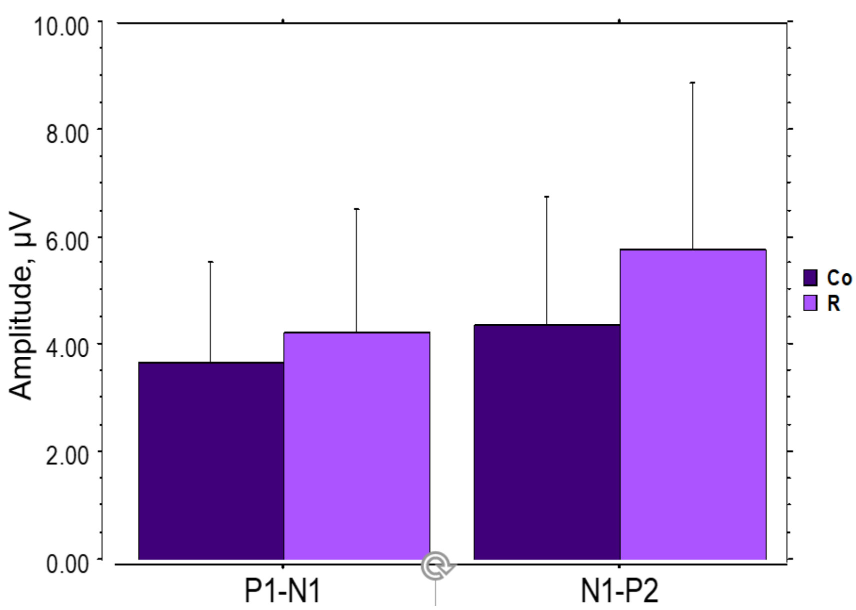


Figure 9. Average CAEP amplitudes in co-located (Co) and spatially separated (R) conditions. The amplitudes were averaged across all SNRs and all levels. Error bars indicate standard deviation.

Speech perception in noise as measured using standardized tests

Performance on the standardized speech perception in noise tests (AzBio and QuickSIN) improved as a function of increasing SNR and spatial separation of noise. Shown in Figure 10, AzBio speech perception scores increased as SNR increased from -5 to 10 dB. Performance significantly benefitted (22% improvement) from spatial separation in the -5 dB SNR condition. QuickSIN performance improved in each level tested as the noise was spatially separated, which is shown in Figure 11. Participants could obtain ≥50% accuracy with a SNR 2-3 dB lower when the noise was spatially separated from the speech (i.e. in spatially separated conditions, participants could maintain accuracy with 2-3 dB more noise than they could in co-located conditions).

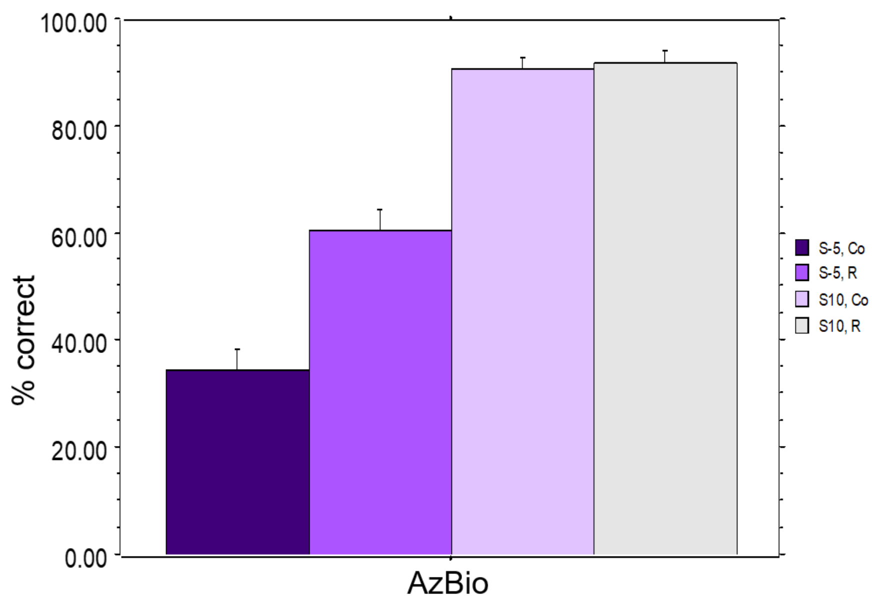


Figure 10. Speech perception scores (% correct) in co-located (Co) and spatially separated (R) conditions as well as -5 dB (S-5) and +10 dB (S10) SNR conditions. 22% improvement in speech perception was observed when the noise was spatially separated from the speech in the -5 dB SNR condition. Spatial separation had little benefit for the +10 SNR condition. Scores improved in both noise location conditions as SNR increased from -5 dB to +10 dB. Error bars indicate standard deviation.

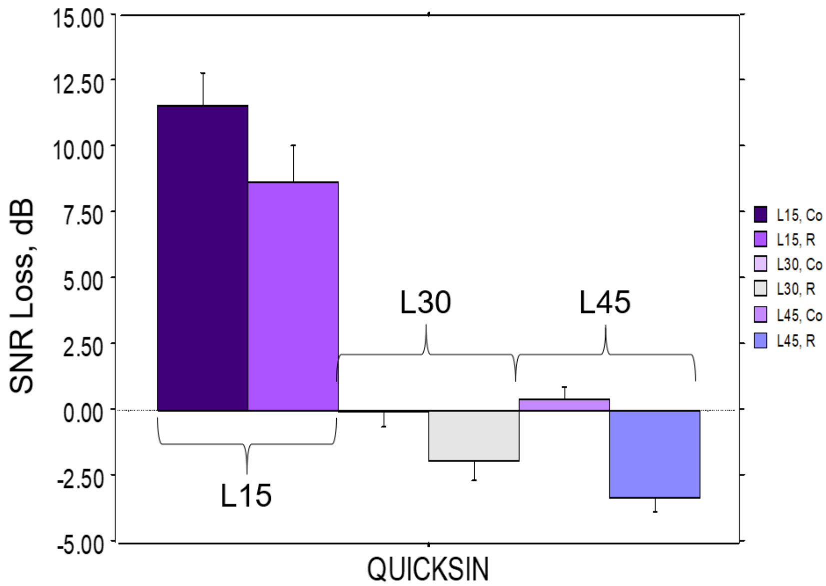


Figure 11. The Y-axis indicates the SNR needed to achieve 50% correct speech perception in noise. Lower SNRs indicate better performance in noise. For each level tested, there was a 2-3 dB improvement in the "SNR Loss" when the noise was spatially separated. Error bars indicate standard deviation.

	Source	**DF**	***F*-value**	***p***
Latency	/DA/ vs /TA/	1,19	6.407	<.05
	Level	3,19	27.802	<.001
	SNR	2,19	52.448	<.001
	Noise location	1,19	39.202	<.001
Amp.	SNR	2,19	10.933	<.002
	Noise location	1,19	15.481	<.002

Table 2. Repeated measures of analysis of variance for latency and amplitude as a function of consonant voicing, level, SNR, and noise location, grouped into the twenty participants rather than grouping each condition individually.

DISCUSSION

The specific aims of this study were to measure how CAEP latencies and amplitudes would be affected by (1) voiced vs. unvoiced consonant-vowel tokens, (2) level, (3) signal-to-noise-ratio (SNR) and (4) co-located vs. spatially separated noise. We also investigated the effect of SNR and noise location on speech perception in noise as measured using standardized tests. There were systematic shifts in latency as a function of consonant voicing, level, SNR and noise location and shifts in amplitude for level, SNR and noise location. Perceptual testing showed an improvement in speech perception in noise as noise was spatially separated, especially in conditions of low SNRs.

Consonant voicing

Voice onset time (VOT) is the time between the onset of the consonant and the start of vocal-fold vibrations, referred to as voicing. Stop consonants /b/, /d/, and /g/ are voiced, and /p/, /t/ and /k/ are unvoiced. Voiced stop consonants such as /d/ are associated with short VOTs, and unvoiced stop consonants such as /t/ with long VOTs. Previous studies have found a positive relationship between CAEP latency and voiced vs unvoiced consonants. Almeqbel (2016) found significant and systematic N2 latency shifts for the different VOTs found in stop CV voiced /ga/ and unvoiced /ka/; the short VOT from /ga/ elicited a shorter N2 latency than the long VOT from /ka/. Sharma and Dorman (1999) used /da/ and /ta/ stimuli to investigate the morphological changes in auditory evoked potentials related to the encoding of VOT and found a systematic shift in N1 latency as VOT increased, suggesting that /da/, with a short VOT, elicits shorter N1 latencies than /ta/, a long VOT stimulus. The systematic latency shifts in P1, N1, and P2 as a function of consonant voicing found in the present study confirm what previous research has found: the difference in the brain's perception of these consonant-vowel tokens are observed through the differences in the CAEP components' latencies. These VOT differences are also evident in infants (Novak et al., 1989).

Level

The present study corroborated the well-known relationship between stimulus level and CAEP latency and amplitude (Hall, 1992; Picton, 2010). As the level of the speech stimulus increased, P1, N1, and P2 latencies decreased and P1-N1 and N2-P2 amplitudes increased. The greatest latency shift is seen between 15 dB SL and 30 dB SL. This is not surprising, because response latency systematically decreases as stimulus intensity level is increased, up until about 40dB or higher, when it then begins to plateau (Hall 1992). The 15 and 30 dB SL levels in the present study were both, on average, equivalent to SPL levels under 40 dB, and showed a decrease in latency as level increased. Moreover, the 45 and 60 dB SL levels were both well above 40 dB SPL (50 and 70 dB SPL, on average), which explains why there were smaller latency shifts between the higher levels in comparison to the lower. Our findings are also congruent with the known underlying mechanisms of neural synchrony and recruitment of nerve fibers in their relation to stimulus level: as level increases, neural synchrony and the recruitment of nerve fibers also increase.

SNR

Previous studies, such as Billings, Tremblay, Stecker and Tolin (2009), have found that SNR has a significant effect on CAEP responses. Billings et al. (2009) recorded CAEP responses to a 1kHz tone presented at 60 and 75 dB SPL in various levels of background noise and measured the changes in latency and amplitude of P1, N1, P2, and N2. They found that amplitude increased and latency decreased with increasing SNR. The data from the present study are congruent with those of Billings et al. (2009).

Noise location

Latencies increased and amplitudes decreased in the co-located noise condition compared to quiet. This agrees with previous studies' findings (Martin et al., 1999; Billings et al., 2009; Whiting et al., 1998) and demonstrates masking a speech token by noise on the electrophysiologic level. When the noise was spatially separated, latencies decreased and amplitudes increased significantly. These amplitude and latency changes with noise location suggest that there is spatial release from masking. The magnitude of the latency and amplitude differences for noise location alone are equivalent to those observed when SNR is improved by 15 dB.

Perceptual tests

The perceptual tests of speech perception in noise also demonstrate the benefits of spatial release from masking. Performance improved as a function spatial separation of the noise, as demonstrated by many studies before (Dirks and Wilson, 1969; Bronkhorst, 2000). The greatest improvement in performance on the AzBio test was when the speech and noise were separated in the -5 dB SNR condition. In contrast, no improvement was observed in the 10 dB SNR condition, as both co-located and separated conditions resulted in high scores (refer to Graph 7). The difference in benefit from spatial separation between the two SNR conditions is reasonable, as an SNR of 10 dB generally makes it easy to hear the target regardless of where the source of the background noise is coming from, whereas an SNR of -5 dB requires the target and the masker to be separated for better target perception. The present study is the first demonstration of how the cortical electrophysiology is congruent with the perceptual benefits of spatially separated noise. Spatial separation of noise resulted in earlier latencies and larger amplitudes, just as it resulted in better perceptual performances.

Head shadow effect

For sound waves travelling to the ears from anywhere other than 0° azimuth, the head becomes an obstacle (i.e it acts as a mass). The head blocks frequencies with wavelengths smaller than its width, and the blocking of these frequencies results in the attenuation of the sound at the further ear (see appendix D). This is referred to as the head shadow effect, and the difference in intensity between the two ears that results is referred to as inter-aural level difference (ILD). The frequencies blocked by the head are those above 1500Hz. ILDs owing to head shadow can be as great as 20dB, depending on the frequency (Gelfand p.235-236).

As we interpret the present study's data, we see the latency and amplitude differences are due to the brain taking advantage of inter-aural level differences for high frequencies caused by head shadow. The head shadow effect is going to attenuate the high frequencies of the masker, and the brain benefits from that. In the co-located condition, all frequencies are being masked via the broadband masker (AzBio 10-talker babble). When the masker is moved to 90° azimuth, its higher frequencies (>1500Hz) are blocked by the head, resulting in a partial unmasking of the speech token the higher frequencies. It is worth noting that phase differences in the low frequencies in the spatially separated condition are less than a millisecond, and therefore do not contribute to the spatial release from masking. The importance of high frequency contributions to cortical evoked potential latency and amplitude was demonstrated by Martin et al. (1999).

Martin et el. (1999) investigated how N1 latency and amplitude was affected by masking at different high-pass frequency cut-offs. They recorded CAEP responses to speech tokens /ba/

and /da/ in seven noise conditions: quiet, and high-pass cut-offs at 4kHz, 2kHz, 1kHz, 500Hz, 250Hz, and broadband noise. Significant latency shifts (increases in latency, that is) were found between the quiet and 4kHz through 500Hz masking conditions, and between the 4kHz-2kHz and 1kHz-500Hz conditions. They found that N1 amplitude decreased as the masker cut-off frequency decreased. Martin et al.'s (1999) study demonstrates the significant effect masking in just the high frequencies has; from the quiet condition to the 2kHz cut-off condition, latency increased by approximately 8ms and amplitude decrease by approximately 1µV. Using the findings from Martin et al. 's (1999) study, we can reasonably speculate that the improving shifts we saw in our study's CAEP component latencies and amplitudes when the masker was moved to 90° azimuth was due to the release from masking in the higher frequencies.

To determine the amount of head shadow effect, we measured the SPL of the noise at 0º and 90º azimuth using probe microphones in a participant's ear canals. The differences in level between the two ears were calculated and plotted on a graph (see Appendix E). The head shadow effect was clearly observed, as the far ear experienced the babble noise at ~14 dB lower than the close ear. With level differences that large, a shift in CAEP component latencies and amplitudes as a result of the head shadow effect is not surprising.

Limitations

Although the present study provides an introduction into investigating the electrophysiologic correlates of spatial release from masking, there are some limitations in our results. First, only one position of noise was tested. Recording data from multiple positions along the azimuth axis would provide us a more comprehensive view on how spatial information is processed. Second, we recorded CAEP during passive listening of the speech token, which does not completely mirror the process of spatial release from masking, as speech perception in noise involves cognitive function (i.e. listening to more meaningful information than CV speech tokens). Lastly, we did not measure perceptual differences in CAEP stimuli, but only correlated the CAEP data with AzBio and QuickSIN data. The mismatch between stimuli used for CAEP recordings and perceptual testing produces a question of whether perceptual differences when noise is spatially separated can be seen at the electrophysiologic level.

Conclusion

The results from normal hearing adults has provided a solid start to understanding the neural encoding and perceptual benefits of spatial release from masking, as well as provided a baseline for investigation of SRM in adults with hearing loss. While it is clear that head shadow and interaural level differences had a significant effect on latency and amplitude improvements when spatially separating the noise, it could still be that the central auditory system is able to squelch background noise via processing of spatial information, and that this capacity is enhanced in more challenging listening environments. Perhaps repeating the study with a different masker, one that only masks frequencies below 1500Hz, would provide us more insight into the role of higher-order processing of spatial information. Future research is warranted with combined electrophysiologic and behavioral measures to more fully understand the underlying mechanisms of speech perception in the presence of noise. We must continue this work so that we can continue to explore ways to improve individuals with hearing loss’ experience in navigating noisy environments.

Appendix A – Example AzBio sentence list and grading sheet

AzBio Sentence Test
List 1
MSTB CD – Track 01
(Left Channel = Speech, Right Channel = Noise)

Sentence	Text	Poss	Score
1	I could hear another conversation through the cordless phone.	9	
2	She relied on him for transportation.	6	
3	He was an ordinary person who did extraordinary things.	9	
4	How long has this been going on?	7	
5	His class was on Saturday.	5	
6	She was entitled to a bit of luxury occasionally.	9	
7	The vacation was cancelled on account of weather.	8	
8	The salon is not open on Mondays.	7	
9	She had a way to justify any of her wrongdoing.	10	
10	I feel sorry for my brother.	6	
11	On numerous occasions they left early.	6	
12	In private she let her hair down.	7	
13	A mother always has something better to do.	8	
14	You should be used to taking money from ladies.	9	
15	Who would lie about cancer for attention?	7	
16	Hang the air freshener from your rearview mirror.	8	
17	You can use your computer to make greeting cards.	9	
18	I guess you know what you're doing.	7	
19	You must live in a gingerbread house!	7	
20	The cat was born with six toes.	7	

Words Correct	
Words Possible	**151**
Percent Correct	

Appendix A: An example of an AzBio list and the grading criteria. The participant's responses are graded on how many words in each sentence they correctly repeat back. For example, if the participant said, "she *relies* on him for transportation" instead of "she *relied* on him for transportation", they would only get a score of 5 out of a possible 6. The total amount of words correct is divided by the total words possible for that list of 20 sentences, and the percent correct is calculated.

Appendix B – Example QuickSIN sentence list and grading sheet

TRACK 21

Practice List A		Score
1. The lake sparkled in the red hot sun.	S/N 25	________
2. Tend the sheep while the dog wanders.	S/N 20	________
3. Take two shares as a fair profit.	S/N 15	________
4. North winds bring colds and fevers.	S/N 10	________
5. A sash of gold silk will trim her dress.	S/N 5	________
6. Fake stones shine but cost little.	S/N 0	________
	TOTAL	________

Appendix B: An example of the types of sentences presented in QuickSIN trials and the grading criteria. It is counted how many underlined words in each sentence the participant correctly repeats back. The total words correct (maximum of 30) is then subtracted by 25.5 to calculate the "SNR loss", which is the signal-to-noise-ratio in which the participant can repeat back at least 50% of the target words. The highest "SNR loss" one can score is -4.5 dB. The more negative the SNR loss, the better the participant can perform in noise.

Appendix C – Example of randomized condition table

Subject number: ____________ Date: ____________ Speech token: ________

Threshold: ____________ Order: +30, +15, +60, +45

Location	Condition	SL level	dB SPL	Noise level (dB SPL)
0° azimuth	Quiet	30		-
0° azimuth	Quiet	15		-
0° azimuth	Quiet	60		-
0° azimuth	Quiet	45		-
Co	10 SNR	30		
Co	0 SNR	30		
Co	-5 SNR	30		
R	-5 SNR	30		
R	0 SNR	30		
R	10 SNR	30		
R	-5 SNR	15		
R	0 SNR	15		
R	10 SNR	15		
Co	10 SNR	15		
Co	0 SNR	15		
Co	-5 SNR	15		
Co	10 SNR	60		
Co	0 SNR	60		
Co	-5 SNR	60		
R	-5 SNR	60		
R	0 SNR	60		
R	10 SNR	60		
R	-5 SNR	45		
R	0 SNR	45		
R	10 SNR	45		
Co	10 SNR	45		
Co	0 SNR	45		
Co	-5 SNR	45		

Appendix C: A table organizing the order of conditions for one participant. After the CAEP recordings were done for the quiet conditions, the levels were randomized per participant. The dB SPL and noise level columns were used as an aid by the experimenter to ensure the levels were adjusted properly per the calibration measurements.

Appendix D – Illustration of head shadow

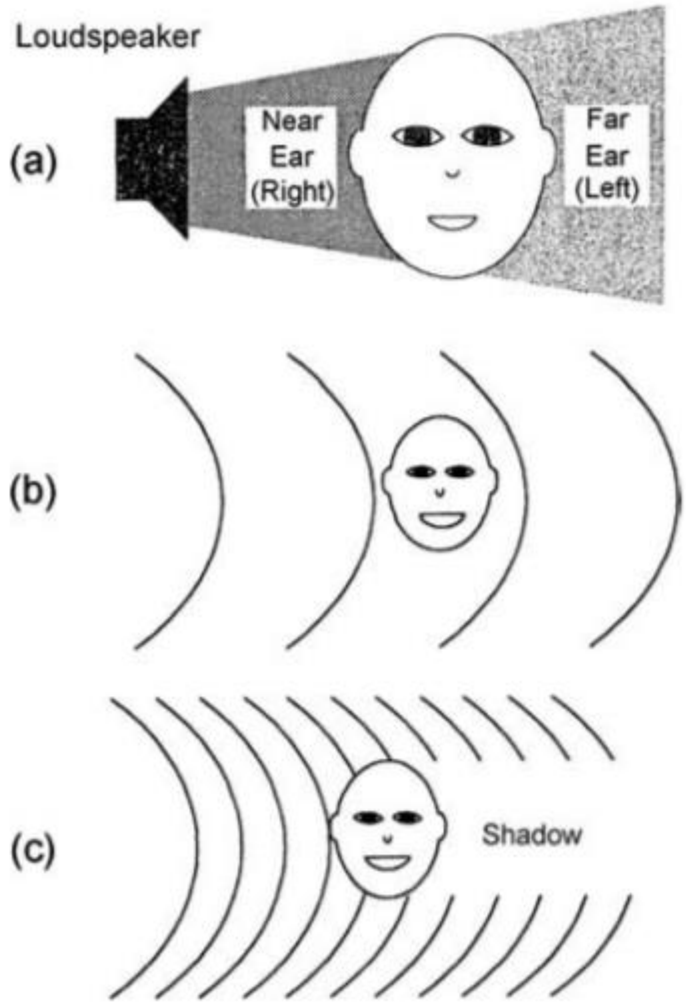
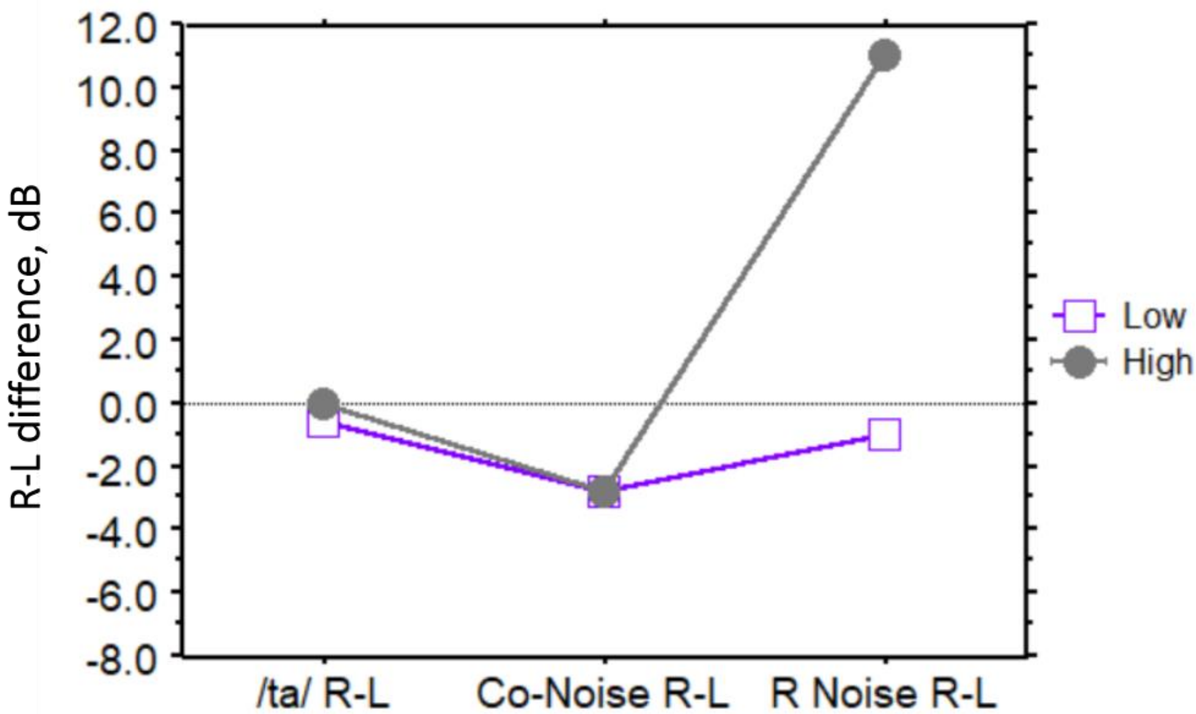


Figure 13.5 (a) Relationship between a loudspeaker and the two ears. (b) Low frequencies bend around the head due to their large wavelengths. (c) High frequencies have wavelengths smaller than head diameter so that an acoustic shadow results at the far ear.

Appendix D: A figure from Gelfand (2009) illustrating head shadow.

Appendix E – Differences in level between the two ears as a function of location



Appendix E: Ear canal probe microphone measure of /ta/ and noise were measured in an individual. The graph plots the difference of right ear minus left ear levels for /ta/ in quiet at 0°, noise masker at 0° (Co-Noise), and at 90° (R Noise), averaged over the low (250-1500 Hz) and high (2000-6000 Hz) frequency ranges. This really highlights the head shadow effect for high frequencies that are on the order of ~14 dB.

REFERENCES

Almeqbel, A. (2016). Temporal encoding of voice onset time in young children. *Otolaryngology: Open Access, 6,* 241-243.

Billings, C. J., Tremblay, K. L., Stecker G. C., & Tolin, W. M. (2009). Human evoked cortical activity to signal-to-noise ratio and absolute signal level. *Hearing Research, 254,* 15-24.

Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: Bradford Books, MIT Press.

Bronkhorst A. (2000). The cocktail party phenomenon: a review of research on speech intelligibility in multiple-talker conditions. *Acta Acustica united with Acustica. 86,* 117–128.

Ching, T. Y. C., van Wanrooy, E., Dillon, H., & Carter, L. (2011). Spatial release from masking in normal-hearing children and children who use hearing aids. *The Journal of the Acoustical Society of America, 129,* 368-375.

Dirks D. D., & Wilson R. H. (1969). The effect of spatially separated sound sources on speech intelligibility. *Journal of Speech, Language, and Hearing Research, 12,* 5–38.

Gelfand, S. A. (2009). Chapter 13: Binaural and Spatial Hearing. In *Hearing: An Introduction to Psychological and Physiological Acoustics* (5th ed., pp. 231-256). Informa Healthcare.

Hall, J. W. (1992). *Handbook of auditory evoked responses.* Allyn and Bacon.

Martin, B. A., Kurtzberg, D., & Stapells, D. R. (1999). The effects of decreased audibility produced by high-pass noise masking on N1 and the mismatch negativity to speech sounds /ba/ and /da/. *Journal of Speech, Language, and Hearing Research, 42,* 271-286.

Novak, G. P., Kurtzberg, D., Kreuzer, J. A., & Vaughan, Jr., H. G. (1989). Cortical responses to speech sounds and their formants in normal infants: maturational sequence and spatiotemporal analysis. *Electroencephalography and clinical Neurophysiology, 73,* 295-305.

Picton, T. W. (2010). *Human auditory evoked potentials*. San Diego, CA: Plural Publishing.

Sharma, A., & Dorman, M. F. (1999). Cortical auditory evoked potential correlates of categorical perception of voice-onset time. *The Journal of the Acoustical Society of America, 106,* 1078-1083.

Whiting, K. A., Martin, B. A., & Stapells, D. R., (1998). The effects of broadband noise masking on cortical event-related potentials to speech sounds /ba/ and /da/. *Ear and Hearing, 19,* 218-231.

Wunderlich, J. L., Cone-Wesson, B. K. (2001). *The Journal of the Acoustical Society of America, 109,* 1526-1537.