

Envelope Frequency Following Responses Are Stronger For High-Pass Than Low-Pass Filtered Vowels

Background: To assess hearing in response to speech, the envelope frequency following response (FFR) can be observed at the fundamental frequency of a vowel stimulus, and its harmonics. FFRs are complex non-linear phenomena, which require better understanding for allowing robust inferences on the assessment of hearing and hearing aid fitting. *Objectives:* To evaluate the effect of stimulus bandwidth on FFR detection rates using filtered vowel stimuli with equal sound levels. *Design:* FFRs were collected whilst presenting repeated vowels (in consonant-vowel-consonant format) filtered into different bandwidths. Eighty stimuli per word were presented at 70 dB SPL LAeq through insert earphones with an inter-stimulus interval of 1s. Responses were detected using frequency-domain Hotelling's T^2 (HT2) tests for individual multiples of the fundamental frequency (F0) and for combinations of F0 multiples. *Study Sample:* Eleven native English-speaking subjects with normal hearing thresholds. *Results:* Average detection rates are highest (69%) with stimuli high-pass filtered >1000 Hz, and significantly lower for low-pass filtered stimuli (40%). *Conclusion:* High-pass filtered vowels elicit stronger FFRs than low-pass filtered vowels at the same dB SPL LAeq. For testing hearing using band-limited speech, filtering effects (due to hearing loss, hearing aid setting or stimulus choice) on responses must be considered.

Keywords: electrophysiology, adult or general hearing screening, speech perception, envelope frequency following responses, objective response detection, Hotelling's T^2 .

Introduction

Current hearing aid fitting algorithms typically prescribe frequency-dependent gains based primarily on pure tone thresholds, after which clinical evaluation of performance is carried out using subjective responses to speech and other sounds (Keidser *et al.* 2012). As subjective responses cannot be easily obtained in some patient groups, such as infants or adults with cognitive impairment, increasing attention has been given to

Filtered word stimuli FFR detection rates

using objective responses for evaluating hearing aids. Auditory brainstem responses (ABRs) to clicks and tone pips and auditory steady state responses (ASSRs) are now routinely used in assessing hearing function in infants (Kennedy *et al.* 2005). However, there are limitations in using these methods for hearing aid fitting evaluation. In particular, very short stimuli can be distorted by hearing aid compression characteristics (Clinard and Tremblay 2013; Gorga *et al.* 1987) and artificial stimuli may be rejected by hearing aid noise reduction algorithms. This has led researchers to explore the use of speech and speech-like stimuli for hearing aid fitting using objective responses. At the brainstem level, these stimuli have included phonemes (Skoe and Kraus 2010), vowels (Choi *et al.* 2013) or word-like stimuli (Easwar *et al.* 2015b) for hearing aid fitting using objective responses. Although beyond the scope of this paper, significant work has also been carried out on measuring responses at cortical levels for hearing aid evaluation (e.g. Purdy *et al.* 2013, Marynewich *et al.* 2012). Brainstem responses to speech-like stimuli can be assessed through envelope frequency following responses (FFRs) (Aiken and Picton 2008). Although comparisons have been made between aided and unaided conditions on FFR detection rate and amplitude (Easwar *et al.* 2015b), there is little data available on how FFRs are affected by filtered stimuli for assessing hearing in specific frequency ranges. In particular, confounding results may arise from higher harmonics eliciting responses at lower frequencies. As most hearing aids are set up using frequency-band specific gains, extending knowledge of the effect of band-pass filtering of speech-like stimuli is vital for further optimising hearing aid fitting and evaluation.

In this work, the effect of filtering stimuli on FFR detection rates is assessed using a Hotelling's T^2 algorithm. The long-term goal of this research is to determine if evoked responses obtained from speech signals can be used to evaluate hearing aid

Filtered word stimuli FFR detection rates

fittings by demonstrating when speech information is reaching the brain. The specific aim in the current work is to assess the effect of vowel bandwidth on FFR response amplitude. The motivation is that by using filtered speech as the stimulus, it may be possible to evaluate the response of the subject in different bands and hence gain information as to how to adjust the frequency response of the hearing aid in that specific frequency band to optimise performance. In previous work (Zhu *et al.* 2013), it was shown that the higher harmonics of a periodic auditory stimulus tend to elicit stronger FFRs at the fundamental frequency, than the lower harmonics of the stimulus. To the authors' knowledge, no previous work has investigated the detection of FFRs to filtered vowel stimuli of equal sound level. In addition, we explore if a multi-frequency test (Hotelling's T^2 for multiple frequencies) can improve response detection sensitivity over single frequency tests.

Materials and Methods

Participants

This study included 11 native English speakers (7 female, 4 male) aged between 20 and 57 years old (33 ± 12 ; mean \pm SD), recruited from staff and students at the University of Southampton. Normal hearing thresholds (≤ 20 dB HL) for frequencies between 500 and 8000 Hz were determined via pure-tone audiometry (PTA), in accordance with the British Society of Audiology (BSA) standards. Otoscopy was also performed to rule out contraindications such as occluding wax, discharge, or foreign bodies in the ear. All participants provided written informed consent to participate. The local ethics committee at the University of Southampton approved the study.

Stimuli

Three stimuli in hVd format (with the vowel V being /ε/, /i/ or /o/, pronounced as ‘hayed’, ‘hid’, ‘hood’, respectively; see also Vanheusden *et al.* (2018)) were presented to the subjects. The stimuli originate from a study by Hillenbrand *et al.* (1995), in which participants were asked to speak slowly and prolong the vowel, which is beneficial for measuring FFRs (Choi *et al.* 2013). The fundamental frequencies (F0) of the stimuli ranged from 161.4-173.9 Hz (ε), 174-180.2 Hz (i) and 165.8-169.8 Hz (o) (high-pitched male speaker). As all vowels are broadband stimuli, stimuli were filtered using a seventh order Butterworth filter to create four band-limited versions of the stimuli: (1) low-pass filtered with a cut-off frequency at 500 Hz, (2) low-pass filtered at 1000 Hz, (3) high-pass filtered at 1000 Hz, and (4) high-pass filtered at 2000 Hz. An example of the spectra for the /ε/ stimulus for each filtered variant is shown in Figure 1A. Given that it is well known that evoked responses become progressively more evident with increasing sound level (Ménard *et al.* 2008), it was decided that the filtered signals should all have the same overall sound pressure level (in dB SPL LAeq). Therefore, all stimuli were presented at 70 dB SPL LAeq via an RME Fireface UC soundcard (Haimhousen, Germany) through ER-2 earphones (Etymotic, IL, USA). Calibration was performed using a Bruel & Kjaer 2260 Investigator and 4157 occluded ear coupler (Royston, Hertfordshire, UK). The inter-stimulus interval was 1s. Stimuli were presented at a sampling rate of 16000 Hz with 16-bit resolution and in both polarities, to reduce stimulus artefacts and linear components of the stimulus responses in the coherent average. Band-limited and wideband stimuli were presented in separate blocks, which were randomized between participants to avoid order effects. Within each block, 80 epochs (40 per polarity) of each vowel were presented randomly. Each block therefore consisted of 240 epochs, with the duration of each epoch being 2 seconds (i.e.

Filtered word stimuli FFR detection rates

a word duration of 1 second and 1 second inter-stimulus interval). Stimulus presentation was controlled using in-house MATLAB scripts (version R2015a, The MathWorks Inc, Natick, MA, USA). Total presentation time was about 40 minutes, with breaks offered between stimulus blocks (or about every 8 minutes).

[Figure 1 near here].

EEG Data Collection and Processing

A 32-channel ActiveTwo system (BioSemi, Amsterdam, The Netherlands) with a standard 10-20 electrode setup was used for collecting EEG data. Two extra electrodes placed at the mastoids acted as reference. Subjects were asked to sit in a reclining chair and were encouraged to sleep during the experiment. Data were collected at 16384 Hz, and response analysis in the current work was performed on data collected from the Cz electrode, as this has been identified as the location with the strongest ABR (Jewett and Williston 1971) and most commercial ASSR systems only allow measurements with a single electrode.

Data processing was performed using in-house software written in MATLAB. Data were band-pass filtered between 80 and 1800 Hz using a seventh order Butterworth filter and afterwards resampled to 4096 Hz. An artefact rejection threshold of 20 μ V was set for removing noisy epochs. The vowel segment for each 1s epoch (vowel durations: / ϵ / 295 ms; /i/ 251.7 ms; / σ / 230.8 ms) was extracted after applying a 10ms delay to account for ABR latency (Easwar *et al.* 2015a). Grand averages for each of the stimuli were generated and their amplitude compared to the amplitude of grand averages generated using a bootstrapping method (Lv *et al.* 2007). Here, random EEG segments not locked to the stimuli were taken from each participant and the average over these segments was calculated. This was repeated 1000 times to create a

Filtered word stimuli FFR detection rates

distribution of amplitudes that might be expected to be obtained from random EEG data. The amplitudes at the 95th percentile of these random EEG averages were then compared to the amplitude of the grand average to determine if a response was present.

Detection of a response for each of the stimuli was determined using a frequency domain Hotelling's T^2 (HT^2) algorithm over the vowel segment excluding the first and last 15 ms of the vowel due to strong fluctuations in the F_0 value (see e.g. Vanheusden *et al.* (2018)). Each epoch was transformed to the frequency domain using a Fast Fourier Transform (FFT), and the HT^2 was applied to the real and imaginary components of the Fast Fourier transformed epochs (Chesnaye *et al.* 2018). This approach was found to outperform alternatives such as the Magnitude Squared Coherence (MSC) and the Fourier Analyser in detecting vowel-evoked FFRs (Vanheusden, 2018). The HT^2 method is similar in concept to the MSC, which detects responses including both the magnitude and phase spectral information of the responses. The HT^2 algorithm also extends the Phase Locking Value (Zhu, 2013), which only uses the phase. Furthermore, the HT^2 algorithm readily lends itself to including multiple frequencies in the same test, as required in the current work. The HT^2 algorithm was first applied to the fundamental frequency F_0 only (denoted by HT^2-1). This was followed by applying the HT^2 algorithm to F_0 and its first two multiples (harmonics; denoted by HT^2-3), and then again to F_0 and its first nine multiples (HT^2-10). The algorithm was also applied to each F_0 multiple separately to evaluate the response at each individual multiple up to 10 x F_0 , with the goal to explore whether specific multiples of F_0 showed stronger responses for different filter settings. It is worth noting that the analysis was limited to the tenth multiple of F_0 , as it has been suggested that neuronal phase locking is limited to about 1500 Hz (Gelfand 2017).

Statistics

Cochran's Q-test was first used on the binary detection data (1 for detection, 0 for no detection) to determine if differences in detection rates over all different types of stimulus filter settings and HT² algorithms (i.e. a comparison of all detection rates over HT²-1, HT²-3 and HT²-10 for unfiltered, low-pass and high-pass filtered stimuli) were significant at an alpha-level of 0.05, irrespective of the vowel stimulus used. The Cochran's Q-test can be considered a variant of the ANOVA test for binary data. The analysis is mainly focused on results from individual subjects rather than the grand average across participants, since future clinical applications require response detection in individuals. As a follow up, a McNemar test with Bonferroni correction was used to determine significant differences in detection rates between individual HT² and stimulus filter settings. As detection times did not show a normal distribution (as shown by the Shapiro-Wilk test), Friedman tests were used to determine if the detection time was significantly different between HT² algorithms and filter settings. Statistical analysis was performed using SPSS Statistics 22 (IBM, Armonk, NY, USA).

Results

A representative coherent average of the EEG for an /ε/ vowel high-pass filtered above 2000 Hz is first shown in Figure 1B, with its power spectrum demonstrating a clear peak at F0 (Figure 1C). Figure 2 provides the grand average responses in the time domain for wideband vowel stimuli /ε/ /i/ and /ʊ/ respectively (black lines) as well as the average responses of bootstrapped data (1000 epochs from random EEG data) with 95% confidence intervals (grey lines). It can be observed that some peaks are above the noise floor, yet a response to F0 is difficult to visualise, indicating the challenge of analysis in the time-domain and the need for a frequency-domain approach.

Filtered word stimuli FFR detection rates

[Figure 2 near here]

Figure 3A shows the detection rates over all stimuli under wideband and filtered settings using the HT²-1, HT²-3 and HT²-10 algorithm. Each of the 12 participants listened to the three vowel stimuli (/ε/ /i/ and /v/) for all filter conditions (full band, low-pass below 500 Hz, low-pass below 1000 Hz, high-pass above 1000 Hz and high-pass above 2000 Hz). As such, the combined observations over all stimuli was 36 observations per stimulus filter setting. Figure 3A therefore shows the results for 36 observations per setting (all stimuli combined), whereas Figure 3B-3D give results for 12 observations (per stimulus). The detection rate was highest (76%) when presenting stimuli high-pass filtered at 2000 Hz and applying a HT² algorithm including the first nine multiples of F0 (F0 till 10 x F0). The lowest detection rate (18%) occurred with low-pass filtered stimuli at 500 Hz and using a single-frequency HT² algorithm (HT²-1). Figures 3B, C and D show results for the individual stimuli /ε/ /i/ and /v/ respectively. Using the Cochran's Q-test, significant differences could also be observed between detection rates for individual stimuli ε (Figure 3B, p=0.009) and v (Figure 3D, p<0.001), but not the i stimulus. For the v stimulus, significantly higher detection rates between pairs of stimuli filtered above 1000 Hz and stimuli filtered below 500 Hz could be observed for the HT²-1 and HT²-3 algorithm using a McNemar's with Bonferroni correction (p=0.028, Figure 3D). Similarly, a significantly higher detection rate was observed for stimuli filtered above 1000 Hz compared to wideband stimuli for the HT²-1 algorithm. (Figure 3D). For the ε stimulus, pairs of tests did not show a significant difference (McNemar's test with Bonferroni correction).

[Figure 3 near here]

Table 1 shows the pairwise comparisons for which significant differences were observed over all stimuli (Figure 3A). Using a Cochran's Q-test with multiple

Filtered word stimuli FFR detection rates

comparisons, a significantly higher detection rate for the HT²-1 algorithm was found for stimuli that are high-pass filtered (> 1000 or > 2000 Hz) compared to broadband stimuli ($p = 0.016$) or stimuli low-pass filtered below <500 Hz ($p=0.001$). For the HT²-3 algorithm, broadband ($p=0.04$) and high-pass filtered stimuli ($p=0.002$ for >1000 Hz, $p=0.017$ for >2000 Hz) had a significantly higher detection rate than stimuli filtered below 500 Hz. Moreover, stimuli filtered above 1000 Hz had a significantly higher detection rate than stimuli filtered below 1000 Hz ($p=0.04$). These results indicate stronger responses with a higher likelihood of FFR detection when high-pass filtered stimuli are used.

[Table 1 near here]

When comparing HT² algorithms, the algorithm using only F0 (HT²-1) had significantly lower detection rates compared to HT²-3 ($p=0.02$) and HT²-10 ($p<0.001$) for broadband stimuli. For low-pass filtered stimuli <500 Hz, the HT²-10 algorithm showed significantly higher detection rates than the HT²-1 algorithm, which uses only F0 ($p=0.02$).

Figure 4 indicates detection rates for single-frequency HT² algorithms using individual multiples (harmonics) up to 10 x F0. Although for most conditions the detection rates for F0 were highest, this was not the case for stimuli low-pass filtered at 500 Hz (6 x F0) and wideband stimuli (2 x F0). Together with Figure 3, these observations suggest that multi-frequency algorithms for detection of FFRs are more sensitive to stimulus responses than algorithms focussing on a single (F0) frequency, as it allows inclusion of F0 multiples which may have a stronger response than F0 itself.

[Figure 4 near here]

Discussion

Objective auditory evoked response detection has become a vital tool for early identification of hearing dysfunction in infants and other patient groups in which behavioural responses are difficult to acquire (Kennedy *et al.* 2005). In verifying hearing aid fitting using electrophysiological approaches rather than behavioural tests, much research has focussed on FFRs to vowels (Aiken and Picton 2006; Ananthakrishnan *et al.* 2016; Choi *et al.* 2013; Clinard and Tremblay 2013). FFRs to vowels and other speech-like stimuli have shown potential in assessing how acoustical features of real-world sound are represented in normal hearing groups (Easwar *et al.* 2015a), as well as in patient groups (Easwar *et al.* 2015b; Russo *et al.* 2008). For hearing aid fitting, it is however also necessary to understand how stimulation in different frequency ranges affects the strength and spectral characteristics of responses, as this provides information about the audibility of the stimulus in these frequency ranges, which could allow for more frequency-specific testing of aided hearing. This might better inform an audiologist on how to optimise hearing aid settings.

In this work, the HT^2 method was chosen for objective detection of FFRs, as this detection algorithm outperforms detection of click ABRs compared to the Fsp, Fmp and q-sample uniform scores test methods (Chesnaye *et al.* 2018). For detection of ASSRs, it performs equally well compared to the F-test, coherence synchrony measures and circular Hotelling's T-test methods (Valdes *et al.* 1997). Furthermore, a recent study on the detection of FFRs to vowels showed that a multi-frequency HT^2 test increased detection rates (albeit non-significantly) in a small cohort of subjects with normal hearing thresholds compared to magnitude squared coherence and F-test methods, and significantly higher detection rates could be observed using a multi-channel HT^2 test (Vanheusden *et al.* 2018). The latter work also showed that the HT^2 algorithm can be

Filtered word stimuli FFR detection rates

easily extrapolated from a single-channel, single-frequency algorithm to a multi-channel, multi-frequency (including F0 harmonics), which is not always trivial for other statistical detection methods (e.g. de Sá *et al.* (2004)).

The current work shows that the most sensitive single-channel HT² methods can detect responses after 80 stimuli in 76% of normal-hearing individuals (>2000 Hz, Figure 3A). Furthermore, high-frequency components of the vowel sounds (>1000 Hz) elicited the strongest responses. Broadband stimulus detection rates are not equal to the sum of detection rates in specific filter bands due to non-linear aspects of speech processing in the auditory pathway and possibly phase difference between the responses to different frequency bands (Zhu *et al.* 2013). Other recent studies have shown that stronger FFRs can be achieved for tone pairs when the envelope phase of the high-frequency tone is delayed to the low-frequency tone (Easwar *et al.* 2018a). Similarly, stronger FFRs could be recorded for some vowels when their high-frequency formants are delayed, indicating the need of adding phase delays for optimal assessment of hearing function to vowel stimuli (Easwar *et al.* 2018b). On the other hand, some studies showed a linear superposition of ASSR amplitudes to modulated tones equalled ASSR amplitudes to individually presented tones (Herdman *et al.* 2002), yet these study did not identify if this effect was due to superposition of independent (brainstem) sources contributing to the ASSR or an effect of neural interactions within a single source of the ASSR (Guérit *et al.* 2017). As is well known for steady state responses in general, the spectral content of the response does not reflect the stimulation frequency directly, but rather the modulation frequency (Aiken and Picton 2008; Greenberg *et al.* 1987). When the spectral content of the stimulus is reduced, this will also affect the modulation and a change in response at the fundamental frequency (and harmonics) may thus be expected. Vowel intensity levels might therefore need to be adjusted to

Filtered word stimuli FFR detection rates

maintain identical sensitivity for different frequency ranges in hearing tests.

Furthermore, a low-frequency hearing loss may not be evident in the response to vowels, when the responses to the higher harmonics in the stimulus are detected; the current results (Fig. 4) confirm that responses to higher harmonics tend to be most evident at the low frequencies (such as F0).

With respect to the detection rates for multi-frequency HT² algorithms on unfiltered vowels, these were similar to the rates observed in previous studies using an F-test after extracting the fundamental frequency using a Fourier analyzer (Choi *et al.* 2013) and a multi-frequency HT² test (Vanheusden *et al.* 2018). The detection rate for the HT²-1 test on F0 (30 %) was lower compared to previous work (Vanheusden *et al.* 2018). Interestingly, detection rates for a HT²-1 algorithm on 2 x F0 did show detection rates as expected from F0 (Figure 4). Reasons for this reduced detection rate at F0 could be a reduced number of stimuli presented in this study compared to other works, which was based on a 75% detection rate in previous work (Vanheusden *et al.* 2018) and a higher fundamental frequency compared to previous work (Choi *et al.* 2013). In addition, the current dataset showed an increased electrical noise level during data collection (Figure 2), and recruited a different (and on average older) test population, which may also have had different physiological noise levels relative to the previous study (Clinard and Tremblay 2013). Overall, results confirm previous findings that including more response features in objective detection algorithms improves test sensitivity (Cebulla *et al.* 2006; Cebulla *et al.* 2001; Picton *et al.* 1987). This is not a foregone conclusion, as including variables that do not show a response, can worsen the results (e.g. de Sá and Felix (2003)).

Over all stimuli, significantly higher detection rates were found for high-pass filtered (>1000 & > 2000 Hz) stimuli compared to stimuli low-pass filtered (<500 Hz)

Filtered word stimuli FFR detection rates

for the HT²-1 and HT²-3 algorithms (Figure 3A). This behaviour could be observed for individual stimuli as well. Significant differences in FFR detection rate were found between high-pass filtered and low-pass filtered versions of the *v* stimulus (Figure 3D). These results show similarities with the findings of Zhu *et al.* (2013), who showed that detection rates to complex tones including low-frequency harmonics of a 100 Hz fundamental frequency had lower detection rates than tones including high-frequency harmonics. In another study, ASSRs in young infants with hearing impairment were also more difficult to obtain for 500 Hz tones compared to higher-frequency tones (Stroebel *et al.* 2007). Earlier work suggested this behaviour might be due to differences in how auditory filter characteristics resolve sound at low-frequency harmonics of the vowel fundamental compared to how filters resolve high-frequency components (Joris *et al.* 2004). Low frequency components (up to 500 Hz) can be resolved well by the auditory periphery, making these components strongly sensitive to stimulus polarity (Moore 2012). Averaging responses of opposite polarity can therefore reduce the strength of the response due to cancellation (Zhu *et al.* 2013). Frequencies above 1000 Hz cannot be resolved well by the auditory periphery. There will however be a stronger response as a higher number of stimulus harmonics falls within the critical band of the high-frequency auditory channels and the degree of phase-locking to the stimulus envelope therefore will increase (Zhu *et al.* 2013). In other words: at higher frequencies, multiple harmonics fall into the same auditory band (as these are wider) and their sum leads to modulated output from these filters (unresolved harmonics, see Laroche *et al.* (2013)). Furthermore, hearing sensitivity is increased in the 1000-4000 Hz frequency range (Moore *et al.* 1997), which can improve response detection, although having stimuli with equal sound level (in dB SPL LAeq) should have accounted for this. In this respect our work provides confirmation that it is not this change in sensitivity (alone)

Filtered word stimuli FFR detection rates

which explains the increased FFR for higher frequency bands. Finally, detection of FFRs to F0 for stimuli with a ‘missing fundamental’ has been observed in several studies (Greenberg *et al.* 1987). Interestingly, increased brainstem response detection rates with increasing frequency appears opposite to the behaviour found for cortical evoked potentials, which increase in amplitude (and therefore likelihood of detection) with decreasing stimulus frequency (Dimitrijevic *et al.* 2008; Evans and Deatherage 1969; Glista *et al.* 2012). This is probably due to differences in characteristics of auditory brainstem and cortical source generators (Billings 2013; Clinard and Tremblay 2013; Gelfand 2017; Korczak *et al.* 2012; Moore 2012).

It should be noted that in the current work all stimuli were presented at the same loudness in terms of dB SPL LA_{eq}. This means that the levels of individual harmonics varied between the stimuli. For example: in the low-pass filtered condition, low frequency harmonics are increased in level compared to the broad-band stimuli to maintain the loudness of the stimulus. This is similar to the approach in Zhu (2013), though they did not specifically use A-weighting. The rationale was that in the longer term we would like to evaluate hearing aids with stimuli of approximately the same loudness at different frequencies. This approach however does not show the relative contribution of low and high frequencies to the FFR with broadband stimuli (as common in real-world listening scenarios). This is an area of future study, where we wish to filter stimuli without maintaining the overall level and to compare the relative importance of different harmonics in the FFR response. The confounding effects of changes in bandwidth and sound level however would still be difficult to disentangle, as in speech sounds different harmonics also have different levels of power. It would still remain unclear if changes in FFRs are due to the changes in bandwidth or changes in overall sound level.

Filtered word stimuli FFR detection rates

For individual vowel stimuli, differences in detection rates could be observed within and between stimuli. The within-stimulus detection rates roughly followed the overall behaviour, with lower detection rates for low-pass filtered stimuli. Between stimuli, the ε and v stimuli showed slightly (but not significantly) higher detection rates to high-pass filtered responses than the i stimulus. For low-pass filtered and wideband stimuli, response detection rates were lowest for v compared to the i and ε stimulus. Differences in detection rates for individual stimuli have been observed in normal hearing (Aiken and Picton 2006; 2008; Choi *et al.* 2013) and hearing-impaired participants (Easwar *et al.* 2015b). These studies also suggested the reason for differences in detection rates are complex, with contributing characteristics being F0 variation, spectral and temporal envelope amplitudes (Choi *et al.* 2013) and harmonic distortions (Aiken and Picton 2008). The latter study also indicated that F0 harmonics close to formant frequencies are more likely to elicit responses for wideband stimuli, which was shown for the ε stimulus in this study (first formant close to $2 \times F0$). Studies on inferior colliculus neuron models have however shown that higher FFR amplitudes are not necessarily to be expected around formant frequencies (Carney *et al.* 2015).

Two main considerations can therefore be made from the current observations. Firstly, it is possible to detect FFRs to vowels filtered into individual frequency bands, although detection rates are better for high-pass filtered relative to low-pass filtered vowels, presented at the same level (dBAeq). Secondly, multi-frequency detection algorithms should be used, as they perform better than single frequency algorithms. The strongest responses tended to occur at lower frequencies, even with the high-pass stimuli. Taken as a whole, it might be possible to use filtered vowel stimuli to assess aided hearing function in different frequency bands by detecting responses using a multi-frequency HT²-10 algorithm. However, the number of stimuli needed would have

Filtered word stimuli FFR detection rates

to be increased above 80 to reach 100% detection in all (normal hearing) subjects.

Findings from this study confirm that FFR to filtered speech stimuli might provide a feasible option for testing a subject's access to speech and evaluating hearing aid fittings.

Further study is nevertheless needed to explore if reliable response detection using such a stimulus paradigm can be achieved for hearing aid users in reasonable clinical time scales. Objective response thresholds are usually higher than behavioural threshold (Billings 2013) and the effect of presenting vowel stimuli in the current format through hearing aids needs to be further evaluated, as hearing aid processing algorithms have for example been shown to sometimes reduce stimulus' output levels when played in isolation (Easwar *et al.* 2012).

Besides understanding the effect of hearing aid processing on test stimuli, further work will be needed to evaluate changes in FFR morphology between aided and unaided conditions for hearing impaired subjects. Such work has been performed for cortical responses (Glista *et al.* 2012; Marynewich *et al.* 2012), but has shown complex effects. Responses may not always be detected under normal hearing conditions (Glista *et al.* 2012) and responses may not always increase in amplitude for aided compared to unaided conditions (Marynewich *et al.* 2012), but see Glista *et al.* (2012)). Clinical application of FFR (and other objective responses to sound or speech stimuli) will only become appropriate once the approaches have been well validated on hearing impaired subjects.

Conclusion

The effect of filtered vowel stimuli on FFR detection rates has been explored with the overall goal to determine how FFRs can be used to assess hearing function in different

frequency ranges. Results show that at equal intensity levels, high-pass filtered vowels show a higher detection rate (i.e. stronger responses) at the fundamental frequency using HT² algorithms compared to low-pass filtered vowels. When applying the HT² algorithm, the use of multi-frequency HT² is recommended for the detection of responses, as the highest detection rates can sometimes be seen for harmonics other than F₀ and significant responses were seen across all harmonics examined (up to about 1500 Hz). There is potential for FFRs from filtered speech to assess hearing within specific frequency ranges.

Acknowledgements, The authors would like to thank Louise Godwin for her technical support in the experimental setup. This research project is funded by the Engineering and Physical Sciences Research Council (EPSRC, grant No. EP/M026728/1). All data supporting this study are openly available from the University of Southampton repository at DOI: <https://doi.org/10.5258/SOTON/D0510>.

References

- Aiken, S.J. and T.W. Picton. 2006. Envelope following responses to natural vowels. *Audiology and Neurotology* 11, no 4: 213-32.
- Aiken, S.J. and T.W. Picton. 2008. Envelope and spectral frequency-following responses to vowel sounds. *Hearing research* 245, no 1-2: 35-47.
- Ananthakrishnan, S., A. Krishnan and E. Bartlett. 2016. Human frequency following response: Neural representation of envelope and temporal fine structure in listeners with normal hearing and sensorineural hearing loss. *Ear and hearing* 37, no 2: e91.
- Billings, C.J. 2013. Uses and limitations of electrophysiology with hearing aids. *Seminars in Hearing* 34, no 4: 257-69.
- Carney, L.H., T. Li and J.M. McDonough. 2015. Speech coding in the brain: Representation of vowel formants by midbrain neurons tuned to sound fluctuations. *eNeuro* 2, no 4: 1-12.
- Cebulla, M., E. Stürzebecher and C. Elberling. 2006. Objective detection of auditory steady-state responses: Comparison of one-sample and q-sample tests. *Journal of the American Academy of Audiology* 17, no 2: 93-103.
- Cebulla, M., E. Stürzebecher and K.-D. Wernecke. 2001. Objective detection of the amplitude modulation following response (amfr): Detención objetiva de la respuesta consecuente de amplitud modulada (amfr). *Audiology* 40, no 5: 245-52.
- Chesnaye, M., S. Bell, J. Harte and D. Simpson. 2018. Objective measures for detecting the auditory brainstem response: Comparisons of specificity, sensitivity and detection time. *International journal of audiology* 57, no 6: 468-78.

- Choi, J.M., D.W. Purcell, J.-a.M. Coyne and S.J. Aiken. 2013. Envelope following responses elicited by english sentences. *Ear and hearing* 34, no 5: 637-50.
- Clinard, C.G. and K.L. Tremblay. 2013. What brainstem recordings may or may not be able to tell us about hearing aid–amplified signals. In *Seminars in Hearing*, 270-77: Thieme Medical Publishers.
- De Sá, A.M.F.M. and L.B. Felix. 2003. Multi-channel evoked response detection using only phase information. *Journal of neuroscience methods* 129, no 1: 1-10.
- De Sá, A.M.F.M., L.B. Felix and A.F.C. Infantosi. 2004. A matrix-based algorithm for estimating multiple coherence of a periodic signal and its application to the multichannel eeg during sensory stimulation. *IEEE Transactions on Biomedical Engineering* 51, no 7: 1140-46.
- Dimitrijevic, A., H.J. Michalewski, F.-G. Zeng, H. Pratt and A. Starr. 2008. Frequency changes in a continuous tone: Auditory cortical potentials. *Clinical Neurophysiology* 119, no 9: 2111-24.
- Easwar, V., D.W. Purcell, S.J. Aiken, V. Parsa and S.D. Scollie. 2015a. Effect of stimulus level and bandwidth on speech-evoked envelope following responses in adults with normal hearing. *Ear and hearing* 36, no 6: 619-34.
- Easwar, V., D.W. Purcell, S.J. Aiken, V. Parsa and S.D. Scollie. 2015b. Evaluation of speech-evoked envelope following responses as an objective aided outcome measure: Effect of stimulus level, bandwidth, and amplification in adults with hearing loss. *Ear and hearing* 36, no 6: 635-52.
- Easwar, V., A. Banyard, S.J. Aiken and D.W. Purcell. 2018a. Phase delays between tone pairs reveal interaction in scalp-recorded envelope following responses. *Neuroscience Letters* 665: 257-262.
- Easwar, V., A. Banyard, S.J. Aiken and D.W. Purcell. 2018b. Phase-locked responses to the vowel envelope vary in scalp-recorded amplitude due to across-frequency response interactions. *European Journal of Neuroscience* 48, no 10: 3126-3145.
- Easwar, V., D.W. Purcell and S.D. Scollie. 2012. Electroacoustic comparison of hearing aid output of phonemes in running speech versus isolation: Implications for aided cortical auditory evoked potentials testing. *International Journal of Otolaryngology* 2012.
- Evans, T.R. and B.H. Deatherage. 1969. The effect of frequency on the auditory evoked response. *Psychonomic Science* 15, no 2: 95-96.
- Gelfand, S.A. 2017. *Hearing: An introduction to psychological and physiological acoustics*: CRC Press.
- Glista, D., V. Easwar, D.W. Purcell and S. Scollie. 2012. A pilot study on cortical auditory evoked potentials in children: Aided caeps reflect improved high-frequency audibility with frequency compression hearing aid technology. *International Journal of Otolaryngology* 2012.
- Gorga, M.P., K.A. Beauchaine and J.K. Reiland. 1987. Comparison of onset and steady-state responses of hearing aids: Implications for use of the auditory brainstem response in the selection of hearing aids. *Journal of Speech, Language, and Hearing Research* 30, no 1: 130-36.
- Greenberg, S., J.T. Marsh, W.S. Brown and J.C. Smith. 1987. Neural temporal coding of low pitch. I. Human frequency-following responses to complex tones. *Hearing research* 25, no 2-3: 91-114.
- Guérit, F., J. Marozeau and B. Epp. 2017. Linear combination of auditory steady-state responses evoked by co-modulated tones. *The Journal of the Acoustical Society of America* 142, no 4: EL395-EL400.

- Herdman, A.T., T.W. Picton and D.R. Stapells. 2002. Place specificity of multiple auditory steady-state responses. *The Journal of the Acoustical Society of America* 112, no 4: 1569-1582.
- Hillenbrand, J., L.A. Getty, M.J. Clark and K. Wheeler. 1995. Acoustic characteristics of american english vowels. *The Journal of the Acoustical Society of America* 97, no 5: 3099-111.
- Jewett, D.L. and J.S. Williston. 1971. Auditory-evoked far fields averaged from the scalp of humans. *Brain* 94, no 4: 681-96.
- Joris, P., C. Schreiner and A. Rees. 2004. Neural processing of amplitude-modulated sounds. *Physiological reviews* 84, no 2: 541-77.
- Keidser, G., H. Dillon, L. Carter and A. O'brien. 2012. Nal-nl2 empirical adjustments. *Trends in amplification* 16, no 4: 211-23.
- Kennedy, C., D. Mccann, M.J. Campbell, L. Kimm and R. Thornton. 2005. Universal newborn screening for permanent childhood hearing impairment: An 8-year follow-up of a controlled trial. *The Lancet* 366, no 9486: 660-62.
- Korczak, P.A., J. Smart, R. Delgado, T. M Strobel and C. Bradford. 2012. Auditory steady-state responses. *Journal of the American Academy of Audiology* 23, no 3: 146-70.
- Laroche, M., H.R. Dajani, F. Prévost and A.M. Marcoux. 2013. Brainstem auditory responses to resolved and unresolved harmonics of a synthetic vowel in quiet and noise. *Ear and hearing* 34, no 1: 63-74.
- Lv, J., D.M. Simpson and S.L. Bell. 2007. Objective detection of evoked potentials using a bootstrap technique. *Medical Engineering and Physics* 29, no 2: 191-98.
- Marynewich, S., L.M. Jenstad and D.R. Stapells. 2012. Slow cortical potentials and amplification—part i: N1-p2 measures. *International Journal of Otolaryngology* 2012.
- Ménard, M., S. Gallègo, C. Berger-Vachon, L. Collet and H. Thai-Van. 2008. Relationship between loudness growth function and auditory steady-state response in normal-hearing subjects. *Hearing research* 235, no 1-2: 105-13.
- Moore, B.C. 2012. *An introduction to the psychology of hearing*: Brill.
- Moore, B.C., B.R. Glasberg and T. Baer. 1997. A model for the prediction of thresholds, loudness, and partial loudness. *Journal of the Audio Engineering Society* 45, no 4: 224-40.
- Picton, T.W., J. Vajsar, R. Rodriguez and K.B. Campbell. 1987. Reliability estimates for steady-state evoked potentials. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section* 68, no 2: 119-31.
- Purdy, S.C., M. Sharma, K.J. Munro and C.L.A Morgan. 2013. Stimulus level effects on speech-evoked obligatory cortical auditory evoked potentials in infants with normal hearing. *Clinical Neurophysiology* 124, no 3: 474-80.
- Russo, N., E. Skoe, B. Trommer, T. Nicol, S. Zecker, A. Bradlow and N. Kraus. 2008. Deficient brainstem encoding of pitch in children with autism spectrum disorders. *Clinical Neurophysiology* 119, no 8: 1720-31.
- Skoe, E. and N. Kraus. 2010. Auditory brainstem response to complex sounds: A tutorial. *Ear and hearing* 31, no 3: 302.
- Stroebel, D., D. Swanepoel and E. Groenewald. 2007. Aided auditory steady-state responses in infants: Respuestas auditivas de estado estable en niños con auxiliares auditivos. *International journal of audiology* 46, no 6: 287-92.
- Valdes, J.L., M.C. Perez-Abalo, V. Martin, G. Savio, C. Sierra, E. Rodriguez and O. Lins. 1997. Comparison of statistical indicators for the automatic detection of 80 hz auditory steady state responses. *Ear and hearing* 18, no 5: 420-29.

Filtered word stimuli FFR detection rates

Vanheusden, F.J., S.L. Bell, M.A. Chesnaye and D.M. Simpson. 2018. Improved detection of vowel envelope frequency following responses using hotelling's t2 analysis. *Ear and hearing*.

Zhu, L., H. Bharadwaj, J. Xia and B. Shinn-Cunningham. 2013. A comparison of spectral magnitude and phase-locking value analyses of the frequency-following response to complex tones. *The Journal of the Acoustical Society of America* 134, no 1: 384-95.

Tables

Table 1. Overview of HT2 algorithm and stimulus filter settings with significant different detection rates over all stimuli (Figure 3A). P-values of Cochran’s Q sample test are given.

Algorithm	Filter Settings	P-Value	Direction of change
HT ² -1 for F0	< 500 Hz vs. > 1000 Hz	0.001	Higher detection > 1000 Hz
	< 500 Hz vs. > 2000 Hz	0.001	Higher detection > 2000 Hz
	Broadband vs. > 1000 Hz	0.016	Higher detection > 1000 Hz
	Broadband vs. > 2000 Hz	0.016	Higher detection > 2000 Hz
HT ² -3	< 500 Hz vs. Broadband	0.04	Higher detection Broadband
	< 500 Hz vs. > 1000 Hz	0.002	Higher detection > 1000 Hz
	< 500 Hz vs. > 2000 Hz	0.017	Higher detection > 2000 Hz
	< 1000 Hz vs. > 1000 Hz	0.04	Higher detection > 1000 Hz
HT ² -1 vs. HT ² -3	Broadband	0.02	Higher detection HT2-3
HT ² -1 vs. HT ² -10	Broadband	<0.001	Higher detection HT2-10
	< 500 Hz	0.02	Higher detection HT2-10

Figure captions

Figure 1. A) Amplitude spectra for the wideband and filtered versions of the / ϵ / stimulus (Amplitude in dB to arbitrary reference). Note that all stimuli have the same power after A-weighting (70 dB SPL LAeq). B) A representative coherent average at EEG electrode Cz to the vowel segment of a /had/ stimulus high-pass filtered above 2000 Hz along with C) its frequency spectrum.

Figure 2: Grand averages for wideband stimuli / ϵ /, /i/ and /v/ respectively (black lines) as well as 95% confidence intervals from random EEG data (grey lines). Random EEG data were generated using a bootstrapping method, in which 1000 random EEG segments were selected from the EEG data of each subject.

Figure 3. Detection rates for stimuli filtered using different bandwidths over A) all three stimuli, B) ϵ , C) i , D) v . P values under the graph title indicate p-values obtained from the Cochran's Q-test. Significant differences between algorithms for the v vowel are given by Mc-Nemar's test with Bonferroni correction. Significant difference in detection rates over all three stimuli (i.e. related to figure 3A) are given in Table 1.

Figure 4. Detection rates for the Hotelling's T2 algorithms when applied to each multiple of F0 A) across three stimuli, B) ϵ , C) I , D) v .