



Brief Report

Objective Detection of the Speech Frequency Following Response (sFFR): A Comparison of Two Methods

Fan-Yin Cheng and Spencer Smith *

Department of Speech, Language and Hearing Sciences, The University of Texas at Austin,
Austin, TX 78712, USA; fanyin.cheng@utexas.edu

* Correspondence: spencer.smith@austin.utexas.edu

Abstract: Speech frequency following responses (sFFRs) are increasingly used in translational auditory research. Statistically-based automated sFFR detection could aid response identification and provide a basis for stopping rules when recording responses in clinical and/or research applications. In this brief report, sFFRs were measured from 18 normal hearing adult listeners in quiet and speech-shaped noise. Two statistically-based automated response detection methods, the F-test and Hotelling's T^2 (HT^2) test, were compared based on detection accuracy and test time. Similar detection accuracy across statistical tests and conditions was observed, although the HT^2 test time was less variable. These findings suggest that automated sFFR detection is robust for responses recorded in quiet and speech-shaped noise using either the F-test or HT^2 test. Future studies evaluating test performance with different stimuli and maskers are warranted to determine if the interchangeability of test performance extends to these conditions.

Keywords: frequency following response (FFR); auditory electrophysiology; Hotelling's T^2 test; F-test; objective detection



Citation: Cheng, F.-Y.; Smith, S. Objective Detection of the Speech Frequency Following Response (sFFR): A Comparison of Two Methods. *Audiol. Res.* **2022**, *12*, 89–94. <https://doi.org/10.3390/audiolres12010010>

Received: 23 December 2021

Accepted: 24 January 2022

Published: 28 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Neural encoding of speech features can be noninvasively assessed using the speech frequency-following response (sFFR). Over the past two decades, the sFFR has been used in a variety of research studies examining: effects of auditory expertise and deprivation on speech feature encoding [1–7]; neural representation of speech in adverse listening conditions [8–10]; objective hearing aid fitting [11]; bimodal hearing assessment [12–14]; and auditory development [15], to name a few examples. Despite the broad use of this technique, few studies, e.g., [11,16–18] have examined statistically-based automated response detection of sFFRs. This consideration is important in both clinical and research domains, as automated detection would aid in objective response identification and would provide a statistical basis for stopping rules when recording sFFRs.

This brief report examined accuracy and test time required to objectively detect sFFRs in quiet and noise using F- and Hotelling's T^2 (HT^2) tests in the frequency domain (see [19,20] for detailed reviews of statistical techniques for objective response detection). The F-test calculates an F ratio of signal power at a single frequency of interest (e.g., the fundamental frequency, F_0) to the mean power of adjacent frequency bins in which no response components are expected. This approach is particularly well-suited for detecting auditory steady state responses in objective audiometry, e.g., [21,22]. The HT^2 test is a multivariate analysis in which mean differences of measured and hypothesized values are tested for significance using multiple sFFR features (e.g., F_0 and harmonics). The HT^2 test is thus ideal for automated sFFR detection, as it analyzes information at multiple frequencies to determine if a response is present [6,23,24]. Further, this test may be more powerful in conditions where individual components are degraded (e.g., when the sFFR is measured in background noise).

The purpose of this study was to objectively detect sFFR responses using both F- and HT² tests and to compare detection accuracies and test times for both approaches. We hypothesized that the HT² test would have better detection accuracy and shorter time-to-detect because it uses more information contained in the sFFR than the F-test.

2. Materials and Methods

2.1. Participants

The University of Texas at Austin approved the methods described in this study. Participants were 18 adults (10 females, age range 20–32). None reported a history of otopathology, neuropathology, or significant noise exposure. All passed otoscopy, tympanometry, and pure tone audiometry screening.

2.2. Stimulus and Recording Procedures

sFFRs were elicited diotically using a 170-ms/da/speech token (F0 = 100 Hz). Stimuli were presented in alternating polarity at 70 dB SPL through electromagnetically shielded ER-3C insert earphones at a rate of 4.3 Hz [13]. In the noise condition, continuous speech-shaped noise was presented diotically at 65 dB SPL. sFFRs were recorded with a Cz-C7 single channel montage for 2000 stimulus repetitions using a Neuroscan SynAmps2 system (Compumedics Neuroscan; Charlotte, NC, USA). Responses were artifact rejected at $\pm 30 \mu\text{V}$ and filtered from 70–2000 Hz using Curry 8 software. Further analyses were conducted offline in MATLAB (The Mathworks, Natick, MA, USA).

2.3. Automatic Response Detection

F-statistic and HT² analyses were performed on frequency-domain data extracted from the 60–180 ms epoch (i.e., steady-state vowel portion) of the “added” sFFR, which biases the responses to reflect speech envelope features (see [25]). The F-statistic approach was implemented on the spectrum of the cumulative average sFFR waveform each time a sweep was added to the average, as described by Picton and colleagues (2003) [20]:

$$F = B \left(x_{F0}^2 + y_{F0}^2 \right) / \sum_{\substack{j = F0 - 5 - B/2 \\ j \neq F0}}^{F0 + 5 + B/2} (x_j^2 + y_j^2)$$

where F0 is the fundamental frequency (averaged across 99–101 Hz bins), j is the frequency of an adjacent noise bin, B is the number of noise bins, and x and y are amplitude and phase components of a specified frequency bin. In the present experiment, $B = 18$, and the upper and lower samples of noise began at $F0 \pm 5$ Hz, respectively. The respective degrees of freedom were 6 and 36 for the numerator and denominator [19,20].

The HT² test was also implemented as described by Picton and colleagues (2003) using the equation [20]:

$$T^2 = N(\bar{x} - \mu_0)S^{-1}(\bar{x} - \mu_0)'$$

where N is the number of sweeps, \bar{x} is a vector of measured dependent variable means with length L , μ_0 is a vector of hypothesized means with length L , and S^{-1} is an inverse covariance matrix of $N \times L$. In the present experiment, both amplitude and phase components from F0, H2, H3, and H4 were used as dependent variables in the HT² test; thus, $L = 8$. Because the data are circular, the vector of hypothesized means (i.e., the expected outcome if a response were not present) consists of zeros. T^2 was converted to an F-statistic using the equation:

$$F = \frac{N - L}{L(N - 1)} T^2 \sim_{F_{df1, df2}}$$

where $df_1 = L$ and $df_2 = N - L$.

The F-statistic produced by each analysis was used to find the statistical probability that an sFFR response was present as a function of sweep count; a schematized overview

of this approach is provided in Figure 1. Detection rate was quantified as the number of sFFRs detected by each statistical test in each listening condition. These proportions were compared using a Chi-square test. We calculated a “time-to-detect” measure as the test time, in seconds, at which detection probability was $\geq 99\%$ and remained above this threshold for at least the following 25 sweeps. A repeated-measures analysis of variance (RMANOVA) was used to assess the effects of two factors (quiet vs. noise; F-test vs. HT² test) on time-to-detect measurements.

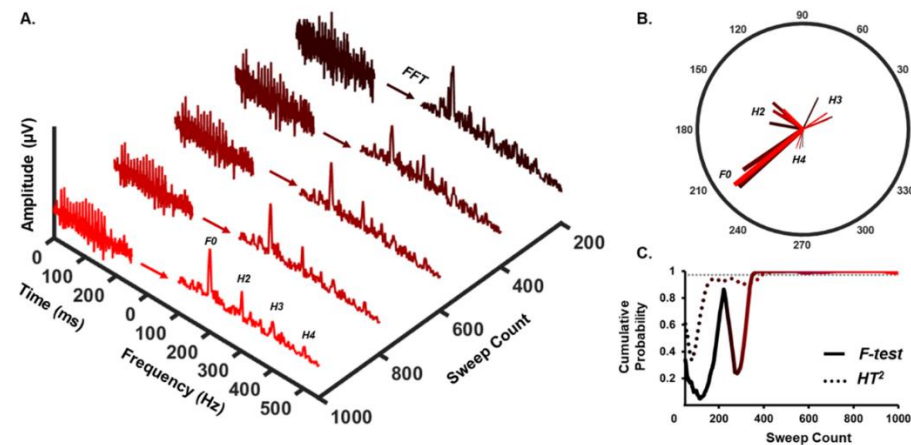


Figure 1. Automatic Response Detection Approach. (A) Average sFFR waveforms collected in quiet (left) and their corresponding spectra (right) are shown as a function of sweep count for a single participant (black-to-red color transition indicates an increase in cumulative sweeps contributing to the average from 200–1000). Frequency bins of interest for the F-statistic (F0 only) and HT² test (F0–H4) are labeled on the final spectrum. (B) Amplitude and phase components are extracted from each frequency bin of interest and used as dependent variables for statistical tests. Note that each vector maintains the color-coding denoting sweep count from panel A. Note also that the F-test only uses magnitude and phase from the F0 vector, whereas HT² uses magnitude and phase components from all vectors shown. (C) Cumulative probability functions are plotted as a function of sweep count for each test with 99% confidence denoted by the gray dotted line. In this single-subject example, F- and HT² tests perform similarly, as an sFFR is detected with >99% confidence between 380–400 sweeps with each test.

3. Results

3.1. Detection Rate

Table 1 lists the percentage of participants from whom sFFR responses could be detected for each condition. Detection rate was similar across conditions [$X(1) = 0.057$, $p = 0.811$], demonstrating that comparable objective sFFR identification was achieved with either statistical approach in quiet or noise.

Table 1. Percentage of response detection across conditions.

	Quiet	Noise
F-statistic	94% (17/18)	100% (18/18)
HT ²	100% (18/18)	94% (17/18)

3.2. Time-to-Detect

Participants in which sFFR responses were not detected in any of the four conditions were removed from the time-to-detect analysis, leaving 16 participants. There was no main effect of statistical test (F-statistic vs. HT² test; $F_{(1,15)} = 1.475$, $p = 0.243$, $\eta_p^2 = 0.09$) or noise condition (quiet vs. noise; $F_{(1,15)} = 0.172$, $p = 0.684$, $\eta_p^2 = 0.011$), nor was there an interaction between factors ($F_{(1,15)} = 3.893$, $p = 0.067$, $\eta_p^2 = 0.206$). As shown in Figure 2, there was a

trend for shorter averaged time-to-detect estimate with the HT^2 test than with F-statistic in quiet, and there is less variability between participants when estimated with HT^2 than F-statistic in both quiet and noise conditions.

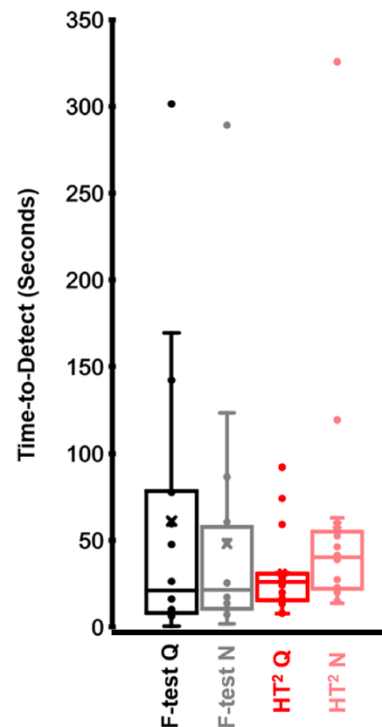


Figure 2. Time-to-detect by statistical test and condition. Each condition is coded with a different color for clarity. The middle line of the box represents the median, and the x represents the mean. The bottom line of the box represents the median of the 1st quartile, and the top line of the box represents the median of the 3rd quartile. The whiskers (vertical lines) extend to minimum and maximum values excluding outliers. Points that exceed 1.5 times of interquartile range are considered outliers. Outliers were not excluded from analyses because they represented the upper limit of test times for our sample.

4. Discussion

We predicted that the HT^2 test, which utilizes more information contained in the sFFR spectrum than the F-test, would have superior detection performance than its counterpart for quiet and noise conditions. While there was a tendency for time-to-detect to be less variable when calculated from the HT^2 test, we found no difference between statistical test performance in quiet or noise on our time-to-detect measure. These findings are broadly consistent with Easwar and colleagues (2020), who found that area under the curve metrics were similar between F- and HT^2 tests. However, the same study indicated that the HT^2 test had higher sensitivity than the F-test. Other reports have also shown time-to-detect advantages of the HT^2 test for sFFRs measured in quiet, e.g., [6]. These discrepancies may be due to differences in sFFR stimuli between studies. For example, Vanheusden and colleagues (2019) used naturalistic monosyllabic words with a dynamic F0 and harmonics to evoke sFFRs, whereas the present study used a synthesized /da/ stimulus with a stable F0 throughout the entire steady-state portion [6]. This likely resulted in robust F0 representation in the sFFR spectra; because this robust component was used in both F- and HT^2 tests, performance of each test may have been more similar than predicted. We also expected that sFFR detection would be superior using the HT^2 test in the noise condition. This prediction was based on the reasoning that noise is deleterious to sFFR spectral peaks, including the F0. Including more spectral peaks in the statistical test would therefore be expected to improve its ability to detect when a “true” sFFR is present. However, because we used speech-shaped noise as a masker, the sFFR components most

impacted by the masker were the harmonics H2-H4 and not the F0. This is because the F0 component of the sFFR is dominated by neural phase-locking from higher frequency channels in which multiple harmonics overlap [18] and may not be appreciably affected by speech-shaped noise, which is low frequency dominant. This again would have made F- and HT² test inputs more similar than expected and may have made their time-to-detect results more comparable in noise. To our knowledge, this is the first report to compare such sFFR detection methods in noise, and future studies with more realistic speech tokens and maskers are warranted.

This brief report adds to the growing literature demonstrating that statistically-based methods for automatic sFFR detection are robust, at least in listeners with normal hearing. Future parametric work is needed to explore the performance of these tests in listeners with hearing loss e.g., [11] or other audiologic disorders as well as the aging e.g., [26]. While objective detection methods are commonly used in commercial devices in which auditory steady-state responses can be obtained (e.g., Interacoustics, Intelligent Hearing Systems, GSI-Audera, Bio-logic), fewer proprietary options exist for sFFR measurement and analysis and many clinician-scientists conduct analyses outside of commercially-available systems (e.g., in MATLAB). This brief report and similar articles provide an evidence base for objective measures that can be employed in commercially available software for sFFR detection and analysis.

Author Contributions: F.-Y.C. and S.S. contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by NIDCD grant K01DC017192.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of The University of Texas at Austin (approval date: 9 January 2018).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data used in this study will be provided without undue reservation upon contacting the senior author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Anderson, S.; White-Schwoch, T.; Choi, H.J.; Kraus, N. Training changes processing of speech cues in older adults with hearing loss. *Front. Syst. Neurosci.* **2013**, *7*, 97. [[CrossRef](#)] [[PubMed](#)]
2. Kraus, N.; Anderson, S.; White-Schwoch, T. The frequency-following response: A window into human communication. In *The Frequency-Following Response*; Springer: Cham, Switzerland, 2017; pp. 1–15.
3. Bidelman, G.M.; Lowther, J.E.; Tak, S.H.; Alain, C. Mild cognitive impairment is characterized by deficient brainstem and cortical representations of speech. *J. Neurosci.* **2017**, *37*, 3610–3620. [[CrossRef](#)] [[PubMed](#)]
4. Krizman, J.; Skoe, E.; Kraus, N. Bilingual enhancements have no socioeconomic boundaries. *Dev. Sci.* **2016**, *19*, 881–891. [[CrossRef](#)] [[PubMed](#)]
5. Krizman, J.; Kraus, N. Analyzing the FFR: A tutorial for decoding the richness of auditory function. *Hear. Res.* **2019**, *382*, 107779. [[CrossRef](#)] [[PubMed](#)]
6. Vanheusden, F.J.; Bell, S.L.; Chesnaye, M.A.; Simpson, D.M. Improved detection of vowel envelope frequency following responses using Hotelling's T2 analysis. *Ear Hear.* **2019**, *40*, 116–127. [[CrossRef](#)] [[PubMed](#)]
7. White-Schwoch, T.; Anderson, S.; Krizman, J.; Nicol, T.; Kraus, N. Case studies in neuroscience: Subcortical origins of the frequency-following response. *J. Neurophysiol.* **2019**, *122*, 844–848. [[CrossRef](#)]
8. McClaskey, C.M.; Dias, J.W.; Harris, K.C. Sustained envelope periodicity representations are associated with speech-in-noise performance in difficult listening conditions for younger and older adults. *J. Neurophysiol.* **2019**, *122*, 1685–1696. [[CrossRef](#)]
9. Smith, S.B.; Cone, B. Efferent unmasking of speech-in-noise encoding? *Int. J. Audiol.* **2021**, *60*, 677–686. [[CrossRef](#)]
10. Yellamsetty, A.; Bidelman, G.M. Brainstem correlates of concurrent speech identification in adverse listening conditions. *Brain Res.* **2019**, *1714*, 182–192. [[CrossRef](#)] [[PubMed](#)]
11. Easwar, V.; Purcell, D.W.; Aiken, S.J.; Parsa, V.; Scollie, S.D. 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 Hear.* **2015**, *36*, 635–652. [[CrossRef](#)]

12. D’Onofrio, K.L.; Caldwell, M.; Limb, C.; Smith, S.; Kessler, D.M.; Gifford, R.H. Musical Emotion Perception in Bimodal Patients: Relative Weighting of Musical Mode and Tempo Cues. *Front. Neurosci.* **2020**, *14*, 114. [[CrossRef](#)] [[PubMed](#)]
13. Kessler, D.M.; Ananthakrishnan, S.; Smith, S.B.; D’Onofrio, K.; Gifford, R.H. Frequency following response and speech recognition benefit for combining a cochlear implant and contralateral hearing aid. *Trends Hear.* **2020**, *24*, 2331216520902001. [[CrossRef](#)] [[PubMed](#)]
14. Xu, C.; Cheng, F.Y.; Medina, S.; Smith, S. Acoustic bandwidth effects on envelope following responses to simulated bimodal hearing. *J. Acoust. Soc. Am.* **2021**, *150*, A64. [[CrossRef](#)]
15. Madrid, A.M.; Walker, K.A.; Smith, S.B.; Hood, L.J.; Prieve, B.A. Relationships between click auditory brainstem response and speech frequency following response with development in infants born preterm. *Hear. Res.* **2021**, *407*, 108277. [[CrossRef](#)] [[PubMed](#)]
16. Easwar, V.; Beamish, L.; Aiken, S.; Choi, J.M.; Scollie, S.; Purcell, D. Sensitivity of envelope following responses to vowel polarity. *Hear. Res.* **2015**, *320*, 38–50. [[CrossRef](#)]
17. Easwar, V.; Birstler, J.; Harrison, A.; Scollie, S.; Purcell, D. The accuracy of envelope following responses in predicting speech audibility. *Ear Hear.* **2020**, *41*, 1732–1746. [[CrossRef](#)]
18. Vanheusden, F.J.; Chesnaye, M.A.; Simpson, D.M.; Bell, S.L. Envelope frequency following responses are stronger for high-pass than low-pass filtered vowels. *Int. J. Audiol.* **2019**, *58*, 355–362. [[CrossRef](#)]
19. Dobie, R.A.; Wilson, M.J. 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.* **1996**, *100*, 2236–2246. [[CrossRef](#)]
20. Picton, T.W.; John, M.S.; Dimitrijevic, A.; Purcell, D. Human auditory steady-state responses: Respuestas auditivas de estado estable en humanos. *Int. J. Audiol.* **2003**, *42*, 177–219. [[CrossRef](#)]
21. Van Maanen, A.; Stapells, D.R. Multiple-ASSR thresholds in infants and young children with hearing loss. *J. Am. Acad. Audiol.* **2010**, *21*, 535–545. [[CrossRef](#)]
22. Rodrigues, G.R.I.; Lewis, D.R. Establishing auditory steady-state response thresholds to narrow band CE-chirps® in full-term neonates. *Int. J. Pediatric Otorhinolaryngol.* **2014**, *78*, 238–243. [[CrossRef](#)] [[PubMed](#)]
23. Hotelling, H. The economics of exhaustible resources. *J. Political Econ.* **1931**, *39*, 137–175. [[CrossRef](#)]
24. Chesnaye, M.A.; Bell, S.L.; Harte, J.M.; Simpson, D.M. Objective measures for detecting the auditory brainstem response: Comparisons of specificity, sensitivity and detection time. *Int. J. Audiol.* **2018**, *57*, 468–478. [[CrossRef](#)] [[PubMed](#)]
25. Skoe, E.; Kraus, N. Auditory brainstem response to complex sounds: A tutorial. *Ear Hear.* **2010**, *31*, 302. [[CrossRef](#)]
26. Clinard, C.G.; Tremblay, K.L. Aging degrades the neural encoding of simple and complex sounds in the human brainstem. *J. Am. Acad. Audiol.* **2013**, *24*, 590–599. [[CrossRef](#)]