A Front-Back Confusion Metric in Horizontal Sound Localization: The FBC Score

Tim Fischer
ARTORG Center for Biomedical Engineering - Hearing Research Laboratory
Bern, Switzerland
tim.fischer@artorg.unibe.ch

Marco Caversaccio
Department of ENT, Head and Neck Surgery, Inselspital, Bern University Hospital
Bern, Switzerland
marco.caversaccio@insel.ch

Wilhelm Wimmer
ARTORG Center for Biomedical Engineering - Hearing Research Laboratory
Bern, Switzerland
wilhelm.wimmer@artorg.unibe.ch

ABSTRACT

In sound localization experiments, currently used metrics for front-back confusion (FBC) analysis weight the occurring FBCs equally, regardless of their deviation from the cone of confusion. To overcome this limitation, we introduce the FBC Score. A sound localization experiment in the horizontal plane with 12 bilaterally implanted cochlear implants (CI) users and 12 normal hearing subjects was performed to validate the method with real data. The overall FBC Rate of the CI users was twice as high as the FBC Score. For the control group, the FBC Rate was 4 times higher than the FBC Score. The results indicate that the FBC Rate is inflated by FBCs that show a considerable deviation from the corresponding value on the cone of confusion.

1 INTRODUCTION

Sound source localization primarily relies on the evaluation of interaural time and level differences [Blauert 1997; Rayleigh 1907]. Front-back confusions (FBCs) occur because sound sources on the cone of confusion are equidistant from the left ear and equidistant from the right ear and thus provide identical interaural time and level differences [Blauert 1997; Rayleigh 1907]. The analysis of FBCs thus provides important information about the extent to which HRTFs or microphone directionality approximate a normal spatial hearing experience.

FBC errors are not simply outliers with extreme error amplitudes. The underlying cause is different from normal localization errors (see Figure 1) which is why FBCs should be analyzed separately [Carlile et al. 1997; Fischer et al. 2020a; Hill et al. 2000; Langendijk et al. 2001; Pastore et al. 2018; Wightman and Kistler 1999]. There is a continuing debate in the literature on the definition of an FBC error. In horizontal sound localization experiments, the most common definition classifies responses crossing the interaural axis as FBCs [Cai et al. 2018; Carlile et al. 1997; Hill et al. 2000; Langendijk et al. 2001; Wightman and Kistler 1999; Zahorik et al. 2006]. This definition, which defines the FBC Rate, is sufficient for experiments that only require a coarse angular resolution of the test setup or feedback method [Montagne and Zhou 2018; Pastore et al. 2018; Wimmer et al. 2017]. However, such setups do not reflect virtual audio scenarios with arbitrary sound source configurations and also limit the measurement resolution for sound localization accuracy.

It has been noted that FBCs occur frequently when sounds are virtually spatialized using individualized, head-related transfer functions (HRTFs) [Morimoto and Aokata 1984; Wightman and Kistler 1989]. By using non-individualized HRTFs, as in virtual reality/audio systems [Iida 2019; Steadman et al. 2019] or hearing aid audio processors [Akeroyd 2014], this negative influence of FBCs on the sound localization ability is further increased [Wenzel et al. 1993]. The analysis of FBCs thus provides important information about the extent to which HRTFs or microphone directionality approximate a normal spatial hearing experience.

FBC errors are not simply outliers with extreme error amplitudes. The underlying cause is different from normal localization errors (see Figure 1) which is why FBCs should be analyzed separately [Carlile et al. 1997; Fischer et al. 2020a; Hill et al. 2000; Langendijk et al. 2001; Pastore et al. 2018; Wightman and Kistler 1999]. There is a continuing debate in the literature on the definition of an FBC error. In horizontal sound localization experiments, the most common definition classifies responses crossing the interaural axis as FBCs [Cai et al. 2018; Carlile et al. 1997; Hill et al. 2000; Langendijk et al. 2001; Wightman and Kistler 1999; Zahorik et al. 2006]. This definition, which defines the FBC Rate, is sufficient for experiments that only require a coarse angular resolution of the test setup or feedback method [Montagne and Zhou 2018; Pastore et al. 2018; Wimmer et al. 2017]. However, such setups do not reflect virtual audio scenarios with arbitrary sound source configurations and also limit the measurement resolution for sound localization accuracy.

A refined FBC definition requires FBC responses to fall within a specific range of a response-dependent regression line, which
The ideal position of an FBC, i.e., the position of a stimulus \( s_i \) mirrored on the interaural axis, is defined as the FBC center \( c_i \). We compute the deviation \( \theta_i \) between the response \( r_i \) and the FBC center \( c_i \) as the shortest absolute angular difference (minor arc). An example for one stimulus-response pair is given in Figure 2. The maximum allowable deviation \( \theta_{\text{max},i} \) is defined by the interaural axis and the FBC center \( c_i \) (see Figure 3). The further away a response \( r_i \) is from its corresponding FBC center and the closer to the interaural axis, the less likely it is to be an FBC. Therefore, we introduce a weighting factor \( w_i \) for each measured FBC:

\[
    w_i = 1 - \frac{\theta_i}{\theta_{\text{max},i}}
\]

where \( \theta_{\text{max},i} \) is the maximum deviation in the direction of the response (clockwise or counterclockwise) under consideration of the interaural axis. The weighting \( w_i \) ranges from 0 (response on the interaural axis) to 1 (a perfect FBC). An illustration of the procedure is given in Figure 3.

To provide a subject-level FBC metric that considers deviations of the responses from the cone of confusion and their proximity to the interaural axis, we propose the FBC Score as defined in Equation 3.

\[
    \text{FBCScore} = \text{FBCRate} \cdot \frac{1}{\hat{N}} \sum_{i=1}^{\hat{N}} w_i
\]

An FBC Score of 100 % indicates that all responses were FBCs and exactly match the stimulus positions mirrored on the interaural axis. In contrast, an FBC Score of 0 % would indicate that the responses did not contain any FBCs during the trial. For a calculation example with numeric data, a step-by-step guide for the calculation of the FBC Score with corresponding explanations is provided in the Appendix A. To facilitate the calculations, we implemented the procedure as a MATLAB function. The function takes the sets of stimulus \( \hat{s} \) and responses \( \hat{r} \) as input parameters. In addition to the calculations described above, a separate analysis of front-back and back-front confusions is performed.

### 3 RESULTS

Table 1 shows the results of the FBC Score analysis compared to the FBC Rate for each CI user and NH subject. For the CI users, the mean overall FBC Score is 27 % in contrast to the almost twice as high mean FBC Rate of 47 % (\( p < .001 \), two-sided Wilcoxon signed rank test). The Wilcoxon signed-rank test was applied because it does not assume normality in the data. It allows to compare two sets of scores, that come from the same participants.

---

1https://www.artorg.unibe.ch/research/hrl/data/fbc_score/
Figure 2: The FBC center $c_1$ is the stimulus origin $s_1$ mirrored on the interaural axis. The interaural axis corresponds to the line from 270° to 90°. Valid FBCs can occur within the light shaded stripe. The deviation of the response $r_1$ from $c_1$ is denoted with $\theta_1$. In this example, the response $r_1$ lies in counter clockwise direction with respect to $c_1$.

Figure 3: Weighting factor calculation for the example shown in Figure 2. The y-axis shows the front-back confusion weighting factor $w$ and the x-axis shows the rear azimuth. The deviation of the response $r_1$ from $c_1$ is denoted by $\theta_1$. The response $r_1$ lies in counter clockwise (cw) direction with respect to $s_1$, therefore $\theta_{\text{max},1}$ is equal to $\theta_{\text{max},1}(\text{cw})$. The areas $\theta_{\text{max},1}(\text{ccw})$ and $\theta_{\text{max},1}(\text{cw})$ are limited by the interaural axis as defined in Figure 2.

Table 1: The table shows the results on subject-level for the front-back confusion (FBC) score analysis compared to the FBC Rate for each cochlear implant (CI) user and normal hearing (NH) subject. Furthermore, an overview of the demographic of the participants is given.

| ID   | FBC Rate (%) | FBC Score (%) | Age (yr.) | CI experience (Left/Right, in yr.) |
|------|--------------|---------------|-----------|-----------------------------------|
| CI01 | 60.0         | 41.3          | 20        | 5/16                              |
| CI02 | 46.7         | 19.8          | 67        | 16/15                             |
| CI03 | 53.3         | 28.8          | 48        | 8/9                               |
| CI04 | 50.0         | 39.9          | 62        | 6/1                               |
| CI05 | 46.7         | 22.7          | 64        | 12/13                             |
| CI06 | 43.3         | 26.9          | 66        | 19/17                             |
| CI07 | 40.0         | 20.1          | 25        | 17/13                             |
| CI08 | 46.7         | 18.2          | 44        | 12/6                              |
| CI09 | 33.3         | 15.4          | 58        | 3/6                               |
| CI10 | 50.0         | 31.3          | 57        | 12/14                             |
| CI11 | 40.0         | 23.4          | 62        | 7/1                               |
| CI12 | 50.0         | 27.8          | 22        | 4/3                               |
| NH01 | 0.0          | 0.0           | 32        | n.a.                              |
| NH02 | 0.0          | 0.0           | 32        | n.a.                              |
| NH03 | 0.0          | 0.0           | 26        | n.a.                              |
| NH04 | 0.0          | 0.0           | 41        | n.a.                              |
| NH05 | 10.0         | 4.2           | 43        | n.a.                              |
| NH06 | 6.7          | 0.0           | 26        | n.a.                              |
| NH07 | 16.7         | 9.4           | 45        | n.a.                              |
| NH08 | 0.0          | 0.0           | 25        | n.a.                              |
| NH09 | 6.7          | 1.3           | 29        | n.a.                              |
| NH10 | 0.0          | 0.0           | 31        | n.a.                              |
| NH11 | 0.0          | 0.0           | 28        | n.a.                              |
| NH12 | 3.3          | 0.0           | 37        | n.a.                              |

For the NH control group, the FBC Rate is 4 times as high as the FBC Score (FBC Score: 1 %, FBC Rate: 4 %; p = 0.063).

4 CONCLUSION AND DISCUSSION

For the evaluation of sound localization experiments, the errors are quantified in terms of their deviation from the stimulus position. However, for FBC analysis, thus far, only a rough distinction between “no FBC” and “FBC” is made, and the occurrences are counted. In this report, we propose a method for quantifying the severity of an FBC, allowing a more specific analysis of this phenomenon.

In the example presented here, the FBC Rate of the CI users would indicate that the tested subjects are prone to FBCs (47 %). However, the FBC Score shows that this assumption does not necessarily hold true since the impact of the FBCs is significantly smaller by a factor of 1.7 (27 %). Therefore, the FBC Rate in our example includes many FBCs with a considerable deviation from the corresponding stimulus position on the cone of confusion, to which this phenomenon actually refers.

A critical examination of the FBC Score could raise the question of whether the calculated score is interpretable and unambiguous. Indeed, an FBC Score value near 0 % could indicate either no FBCs or many FBCs combined with wide scatter of lateral angle judgements. In both scenarios, the FBC Score must ensure that FBCs play no role in the analysis of the subject’s spatial hearing characteristics. The first scenario is trivial, the value equals the result obtained with the FBC Rate. The assumption for the second scenario would be that all localization responses of the subject are found at a very small distance from the interaural axis. An extreme example for this the FBC Score would be, that all stimuli are played frontally from 0° and answered by the subject constantly at 90.1°. This scenario would result in an FBC Rate of 100 %. The FBC Rate would also be 100 % for right-left confusions [Majdak et al. 2011], for example when all stimuli are played from 90.1° position and the subject answers constantly at 270.1° position. In both scenarios, the FBC Rate is deceptive, since the underlying errors are most likely not 100 % due to FBCs. The FBC Score intercepts the fallacy of the FBC Rate according to its calculation rule and shows a slightly higher value than 0 %.
In contrast to a subject who is not susceptible to the FBC error type (FBC Score = 0 \%), an FBC Score value near 100 \% shows the highest possible susceptibility of a subject to FBCs. All responses of the subjects were FBCs and exactly match the stimulus positions mirrored on the interaural axis. As illustrated above, the FBC Rate does not reflect the underlying cause of FBCs as accurate as the FBC Score, because an FBC Rate value of 100 \% does not necessarily indicate a high susceptibility of the subject to FBCs. For any value of the FBC Score, the following applies: FBC Score ≤ FBC Rate. Compared to the FBC Rate, the FBC Score is not so strongly affected by error types that may not be due to FBCs.

In the above section we have shown that although the emergence of an FBC Score value is multicausal, the metric does serve its purpose. It performs a weighting of localization errors that allows to assess the subject’s susceptibility to the FBC error type more specifically than with the FBC Rate. An alternative option to illustrate the subjects’ susceptibility on the phenomenon of FBCs would be to indicate, in addition to the FBC Rate, the localization error with respect to the stimulus’ position mirrored on the interaural axis. In our opinion, however, this approach may overestimate the importance of FBCs caused by stimuli near the interaural axis for two reasons. First, the influence of FBCs near the interaural axis is rather small in real life situations. Second, both error types, i.e., the FBC and localization errors, overlap for stimuli originating from close to the interaural axis, which can lead to distortions in the FBC analysis. By weighting the distance between the stimulus and the interaural axis and by considering the proximity of the response to the interaural axis, this influence on the assessment of the FBC phenomenon can be mitigated with the FBC Score. The effect of this weighting is particularly demonstrated by the 4 times lower FBC Score compared to the FBC Rate in the NH control group. Here, all FBC-causing stimuli originated from the measurement position with the smallest distance to the interaural axis.

The studies in [Macpherson and Middlebrooks 2000; Montagne and Zhou 2018] defined responses as FBCs if the maximum deviation from the FBC center was less than or equal to a 40 or 45-degree angle. Such a static threshold value is not feasible in virtual environments with non-individualized HRTFs or in clinical studies with hearing-impaired subjects, due to the large differences in the inter-subject localization accuracy [Akeroyd 2014; Steadman et al. 2019]. An explanation of the underlying theoretical and psychoacoustic mechanism would be missing for the justification of such a threshold value. Besides the static threshold value, the iterative regression method applied in [Macpherson and Middlebrooks 2000; Montagne and Zhou 2018] relies on the distribution and number of sound localization responses. The FBC Score does neither depend on a predefined threshold value nor on the quality or quantity of the localization results. This may be especially important when evaluating the performance of hearing-impaired subjects in clinical trials. These localization results often show a wide individual variability [Akeroyd 2014] and consist of limited samples due to the limited testing time and subject attention. In addition to evaluate the performance of hearing-impaired subjects, the FBC Score may also be used to evaluate the quality of sound spatialization of a virtual reality system [Steadman et al. 2019].

ACKNOWLEDGMENTS
The authors would like to thank Christoph Schmid from the Department of ENT, Head and Neck Surgery, Bern University Hospital (Inselspital), for the recruitment of the CI users of this study.

REFERENCES
Michael A Akeroyd. 2014. An overview of the major phenomena of the localization of sound sources by normal-hearing, hearing-impaired, and aided listeners. Trends Hear. 18 (12 2014).
Jena. Blauert. 1997. Spatial hearing : the psychophysics of human sound localization. MIT Press, Cambridge, MA.
Yuexin Cai, Guisheng Chen, Xiaoli Zhang, Guangzheng Yu, Hanjie Mo, Jiajia Jiang, Xiaoting Chen, Fei Zhao, and Ying Zheng. 2018. Influence of Audiospatial Training on Horizontal Sound Localization and Its Related ERP Response. Front Hum Neurosci 12 (10 2018), 423.
Simon Carlile, Philip Leong, and Stephanie Hyams. 1997. The nature and distribution of errors in sound localization by human listeners. Hear Res 114, 1-2 (12 1997), 179–196.
Tim Fischer, Martin Kompis, Georgios Mantokoudis, Marco Caversaccio, and Wilhelm Wimmer. 2020a. Dynamic sound field audiometry: Static and dynamic spatial hearing tests in the full horizontal plane. Applied Acoustics 166 (2020), 107363. https://doi.org/10.1016/j.apacoust.2020.107363.
Tim Fischer, Christoph Schmid, Martin Kompis, Georgios Mantokoudis, Marco Caversaccio, and Wilhelm Wimmer. 2020b. Pinna-Imitating Microphone Directionality Improves Sound Localization and Discrimination in Bilateral Cochlear Implant Users. Ear Hearing (in print) (2020).
P. A. Hill, P. A. Nelson, O. Kirkeby, and H. Hamada. 2000. Resolution of front–back confusion in virtual acoustic imaging systems. J Acoust Soc Am 108, 6 (12 2000), 2901–2910.
Kazuhiro Iida. 2019. Head-Related Transfer Function and Acoustic Virtual Reality. Springer. 30 pages.
Erno H. A. Langendijk, Doris J. Kistler, and Frederic L. Wightman. 2001. Sound localization in the presence of one or two distracters. J Acoust Soc Am 109, 5 (5 2001), 2123–2134.
Tomasz Letowski and Szymon Letowski. 2011. Localization Error: Accuracy and Precision of Auditory Localization. In Advances in Sound Localization, Pawel Strumillo (Ed.). InterelOpen, Rijeka, Croatia. 4 https://doi.org/10.5772/15652.
Ewan A. Macpherson and John C. Middlebrooks. 2000. Localization of brief sounds: Effects of level and background noise. J Acoust Soc Am 108, 4 (10 2000), 1834–1849.
Piotr Majdak, Matthew J Goupell, and Bernhard Laback. 2011. Two-dimensional localization of virtual sound sources in cochlear-implant listeners. Ear Hear. 32, 3 (2011), 198–208.
Christopher Montagne and Yi Zhou. 2018. Audiovisual Interactions in Front and Rear Space. Front Psychol 9 (5 2018).
Masayuki Morimoto and Hitoshi Aokata. 1984. Localization cues of sound sources in the upper hemisphere. J Acoust Soc Jpn 5, 3 (1984), 165–173.
M. Torben Pastore, Sarah J. Natale, William A. Yost, and MF. Dormann. 2018. Head Move-ments Allow Listeners Bilaterally Implanted With Cochlear Implants to Resolve Front-Back Confusions. Ear Hear. 39, 6 (2018), 1224–1231.
Lord Rayleigh. 1907. XII. On our perception of sound direction. Lond Edinb Phil Mag 13, 74 (1907), 214–232. https://doi.org/10.1080/147864079463599.
Mark A Steadman, Chungeun Kim, Jean-Hugues Lestang, Dan FM Goodman, and Lorenzo Picinali. 2019. Short-term effects of sound localization training in virtual reality. Sci Rep 9, 1 (2019), 1–17.
Hans Wallach. 1938. On Sound Localization. J Acoust Soc Am 10, 1 (7 1938), 83–83.
Elizabeth M Wenzel, Marianne Arruda, Doris J Kistler, and Frederic L Wightman. 1993. Localization using nonindividualized head-related transfer functions. J Acoust Soc Am 94, 1 (1993), 111–123.
Frederic L Wightman and Doris J Kistler. 1989. Headphone simulation of free-field listening: I. stimulus synthesis. J Acoust Soc Am 85, 2 (1989), 858–867.
Frederic L Wightman and Doris J Kistler. 1999. Resolution of front–back ambiguity in spatial hearing by listener and source movement. J Acoust Soc Am 105, 5 (5 1999), 2841–2853.
Wilhelm Wimmer, Martin Kompis, Christof Stieger, Marco Caversaccio, and Stefan Weder. 2017. Directional microphone contralateral routing of signals in cochlear implant users: A within-subjects comparison. Ear Hear. 38, 3 (2017), 368–373.
Pavel Zahorik, Philbert Bangayavan, V. Sundareswaran, Kenneth Wang, and Clement Tam. 2006. Perceptual recalibration in human sound localization: Learning to remediate front-back reversals. J Acoust Soc Am 120, 1 (7 2006), 343–359.
A Front-Back Confusion Metric in Horizontal Sound Localization: The FBC Score

SAP ’20, September 12–13, 2020, Virtual Event, USA

Table 2: Data of the calculation example.

| Test item | Stimulus (°) | Response (°) | FBC center (°) | Deviation (°) | Maximum (°) | Weighting factor w_i |
|-----------|--------------|--------------|----------------|---------------|-------------|----------------------|
| 1         | 20           | 200          | 160            | 40            | 110         | 0.64                 |
| 2         | 280          | 260          | 260            | 0             | 10          | 1.00                 |
| 3         | 65           | 95           | 115            | 20            | 25          | 0.20                 |
| 4         | 350          | 120          | 190            | 70            | 100         | 0.30                 |

A EXAMPLE CALCULATION OF THE FBC SCORE WITH NUMERIC DATA

In the following, we illustrate the calculation of the FBC Score using a small example data set with 5 stimuli \( \hat{s} \) and responses \( \hat{r} \) from a sound localization test in the horizontal plane (see Figure 4). Table 2 summarizes the results for the computational steps involved.

![Figure 4: Visualization of the data for the example calculation of the front-back confusion (FBC) Score. Stimuli are indicated with \( \hat{s} \) and responses with \( \hat{r} \). In this example, all stimuli lie inside the frontal azimuth and thus responses inside rear azimuth (light shaded area) are considered FBCs.](image)

\( \hat{s} = \{ \hat{s}_1 = 20°, \hat{s}_2 = 280°, \hat{s}_3 = 65°, \hat{s}_4 = 350°, \hat{s}_5 = 315° \} \)

\( \hat{r} = \{ \hat{r}_1 = 200°, \hat{r}_2 = 260°, \hat{r}_3 = 95°, \hat{r}_4 = 120°, \hat{r}_5 = 40° \} \)

Step 1 - Calculate the FBC Rate:

First, the FBC Rate is calculated. It is defined as the rate of the number of responses crossing the interaural axis with respect to the number of presented stimuli (see Eq. 1). Only \( \hat{r}_5 \) does not cross the interaural axis. Excluding \( \hat{s}_5 \) and \( \hat{r}_5 \) results in \( N = 4 \) pairs of stimuli \( s_i \) and responses \( r_i \) and an FBC Rate of 80%.

Step 2 - Calculation of \( c_i \) and \( \theta_i \):

The FBC center \( c_i \) is obtained by mirroring the stimulus position \( s_i \) on the interaural axis. The deviation \( \theta_i \) is calculated as the absolute difference between the response \( r_i \) and \( c_i \) measured over the minor arc. For example, for stimulus 3, we have \( \theta_3 = |r_3 - c_3| = |95° - 115°| = 20° \).

Step 3 - Calculation of \( \theta_{\text{max},i} \):

The maximum deviation \( \theta_{\text{max},i} \) between \( r_i \) and \( c_i \) is limited by the interaural axis. If \( r_i \) lies in clockwise direction of \( c_i \), the clockwise limit is applied \( (\theta_{\text{max},i} = \theta_{\text{max},i}^{(+)}). \) The same applies for the counter clockwise direction.

For example, with stimulus \( s_3 \), the minor arc for the response \( r_3 \) lies in counter clockwise direction to the FBC center \( c_3 \), therefore the maximum deviation is \( \theta_{\text{max},3} = \theta_{\text{max},3}^{(-)} = 25° \).

Step 4 - Calculation of \( w_i \):

With \( \theta_i \) and \( \theta_{\text{max},i} \) we compute the weighting factor \( w_i \) for each stimulus using equation (2).

Step 5 - FBC Score:

We now calculate the FBC Score as defined in Equation (3):

\[
\text{FBC Score} = 80\% \cdot \frac{0.64 + 1.00 + 0.20 + 0.30}{4} = 42.8\%
\]

To facilitate the calculations, we implemented the procedure as a MATLAB function.