# Auditory Motion Perception Documentation

***Release 0.5.0***

**Matthieu Berjon**

**Jan 05, 2019**

# Contents

Auditory motion perception is of part of the audio field that remains quite unknown.

This report deals with several aspects of auditory motion perception during head movements.

Rotating the head in front of a static sound creates dynamic changes in localisation cues that could be mistaken for a source that moves. To interpret these cues correctly, the listener must take the motion of the head into account. Geometrically, the angular velocity of a sound source in the world $(S)$ is the sum of the velocity of head rotation $(H)$ and the angular velocity of the source in the acoustic image $(A) : S = A + H$. Perceived auditory motion is therefore determined by how well the auditory system estimates A and H. We used a psychophysical motion-nulling technique in which the lateral motion of a source was adjusted to determine the velocity at which it appeared stationary during head rotation. If S is recovered veridically, then null velocity should be 0.

Moving sounds were created using a cross-fading technique in which a white noise source was moved across a circular array of speakers by sweeping a spatial Gaussian weighting function. On each trial, a pursuit target swept left then right (or vice versa) followed by a moving test sound. Listeners tracked the pursuit target with their head as accurately as possible, and continued to do so unaccompanied during a third sweep in which the test source was presented. Six observers indicated whether the test source appeared to move left or right across the speakers. By varying the velocity of the test source according to a method of constant stimuli, the null point was estimated from the point of subjective equality of the psychometric function using Probit analysis. Pursuit target speeds of 20, 40, 60 deg/s were investigated. The duration and mean location of the test were randomised across trials to encourage judgements of velocity. Head velocity was recorded.

For all observers, the test sound had to move in the same direction but slower than the head rotation to appear stationary. Because the ability to track the pursuit target varied across observers, data were analysed on the basis of actual head rotation rather than target velocity. This revealed an approximately linear trend with a slope of 0.56. Thus, the test sound had to move around half the speed of the measured head rotation to achieve the null.

The results indicate that perceived motion during head rotation is not veridical; a stationary sound appears to move in the opposite direction to the head movement. H is therefore underestimated with respect to A. The result is similar to that obtained in vision.

**Todo**

This needs to be shorter improved and less based on the introduction.

CHAPTER 1

# Background on hearing

## Auditory localisation

The auditory system is able, even in the absence of visual cues to derive a representation of the world thanks to its two sensors, the ears. The location of a sound source relative to the listener's head can be described in terms of azimuth, elevation and distance.

Localisation in azimuth is mainly attributed to a binaural processing of acoustic cues based on time and intensity differences between the ears. Localisation in elevation is explained by the use of monaural cues although these cues also play a role in azimuth localisation. Localisation in distance can be determined from various acoustic cues related to the transmission of sound over distance, such as intensity and spectral content, and to the effects of acoustic reflections, such as interaural coherence and reverberant tails.

### Localisation in azimuth

#### The duplex theory

*[Ray76]* attempted to account for localisation in azimuth in terms of interaural difference cues. He appreciated that when a sound is presented from the side, the listener's head interrupts the path from the source to the opposite ear. The result is a difference in pressure between the closest ear (ipsilateral) and the farthest ear (contralateral) known as *ILD*. The relative difference will increase with frequency. However, for sources below $1000$ Hz, because the sound wavelength will be several times larger than the head, the head does not present a significant obstacle. Rayleigh pointed out that at these low frequencies, the *ILD* between the two ears would consequently be too small to be perceptible. *[Ray07]* demonstrated that humans are also sensitive to the *ITD* of low frequency pure tones. *ITD* reflect the difference in path distance to each ear when sound source is located to one side (see Fig. 1.1.1). However Rayleigh pointed out that, for a pure tone, the azimuth corresponding to a given *ITD* is ambiguous if the tone's wavelength is less than the width of the of the head. For pure tones, therefore, *ITD* are effective for frequencies whose wavelengths are well below about $20\ cm$, whereas, *ILD* are effective for frequencies whose wavelengths are well above $20\ cm$. This two-mechanism account of sound localisation in azimuth became known as the duplex theory.

Fig. 1.1: Representation of the binaural cues *ITD* and *ILD* (After *[Dan11]*). Representation of Interaural time and intensity differences for a monochromatic sound sound.

The duplex theory was supported by several studies such as *[SN36]*, who found a minimum in accuracy for pure-tone localisation at around 3000 $Hz$ ($\sim$ 11 $cm$). *[STFJ55]* found a minimum around 1500 $Hz$ ($\sim$ 23 $cm$). These results suggest that there is a range between 1500 and 3000$Hz$ where the wavelength is too high to provide adequate *ILD*. At low frequencies, where the wavelength is important compared to the radius head, the sound wave is reflected to a negligible degree, meaning that the *ILD* will be close to 0. It should be noted, however, that this is only true for a source beyond about 1 $m$ where the wave can be treated as planar. Close to the head, the wave front will be spherical and thus subject to the inverse-square relationship between sound intensity and distance, which will have the same effect at all frequencies. The difference in path distance to each ear can thus result in a significant difference in intensity between the two ears, even if no head shadowing occurs *[SCSK00]*.

### Limitation of the binaural cues

*ITD* and *ILD* depend on both frequency and elevation. *[Wal39]* described a form of geometrical locus which has the shape of a cone centred on the interaural axis and corresponding to an infinite number of positions for which the *ITD* and *ILD* are roughly constant. This locus id known as the "cone of confusion" *[WS54]* (see Fig. 1.2). Because many positions on these cones surfaces can correspond to the same pairing of *ITD* and *ILD*, ambiguities in localisation occur, even within the horizontal plane, resulting in front/back errors. *[You31]* showed that head movement can compensate for the lack of pinnae in localisation. This was confirmed later on by *[FF68]* who used a broadband noise pulse and subjects were asked the position of the source according to several conditions such as head restrained of free and with their own pinnae, an artificial pinnae or no pinnae. His finding was that head movements brought in all conditions a very good disambiguation of the source position. *[Wal40]* introduced a general description of the nature of head movements during localisation tasks and pointed out the need for dynamic cues for localisation disambiguation. This was confirmed by *[Bur58]* who compared front/back errors with clamped or free head and with or without covering one or two ears using a noise (per octave band). His conclusions were that disambiguation was almost complete when the head was free. The disambiguation slightly decreased when using noise between 800 and 2400 $Hz$ and decreased dramatically at higher frequencies (above 2400 $Hz$) when both ears are coveredfootnote{The ear away from the loudspeaker was covered with an earphone, which was fed with a wide band random noise in order to mask it at all frequencies.}.

Fig. 1.2: The cone of confusion. Identical values of *ILD* and *ITD* of two opposite points anywhere on the surface of the cone represented by the hyperbolia in two dimensions (After *[Bla83]*).

## Localisation in elevation

The presented localisation cues, based on interaural differences are not sufficient to explain discrimination within the cones of confusion when the head is stationary. *[Ray76]* suggested that spectral cues may play a role. He later confirmed that distorting the acoustics of the pinna (by adding "little reflective flaps") could adversely affect accuracy of front/back judgements (*[Ray07]*). Monaural cues (or spectral cues) can be used to explain discrimination of elevation because the sound is spectrally distorted by reflections and diffractions around the torso, shoulders, head and pinnae before reaching the ear in a way that is dependent on elevation. The resulting colorations for each ear of the source spectra, depending on both direction and frequency, provide a localisation cue. *[LB02]* showed that spectral cue has an impact in localisation in high frequencies and especially, by testing narrow band noises, they suggested that up-down localisation depend upon frequencies between 4 and 16 $kHz$ and front-back localisation on frequencies between 8 and 16 $kHz$. In case of remaining confusion about a source position, *[WK99]* showed that head movements will solve these ambiguities and support the Wallach's theory (*[Wal40][TR67]*).

Fig. 1.3: Representation of main auditory cues used for localisation according to the frequency.

## Localisation in distance

According to *[Rum12]*, there is mainly $4$ cues in localisation in distance:

- the inverse-square law of intensity.
- direct to reverberant ratio.
- small path differences between direct sound and reflections.
- high frequencies attenuation.

### Intensity

In the earliest studies, intensity was considered the primary acoustic cue to distance (*[Tho92]*). *[Edw55]* in two experiments using a metronome and the ticking of a clock. He measured that the *JND* in distance was about $20\ \%$ of overall distance. For a stationary sound source in acoustic free field and emitting uniform spherical waves, the sound source intensity is related to distance from the sound source by an inverse square law. The intensity is related the distance $R$, from the source to the listener by a factor $\frac{1}{R^2}$. Since sound pressure is proportional to the square root of intensity, pressure obeys a $\frac{1}{R}$ relation.

### Reverberation

In any environments with sound reflecting surfaces, the ratio of energy reaching a listener directly to energy reaching a listener after reflecting the surface contact varies systematically with distance. This cue is called the direct-to-reverberant energy ratio and decreases as distance between the listener and source increases. In rooms, change in direct-to-reverberant energy ratio is primarily due to the effect of the inverse-square law on the direct sound because the energy in the later part (all the reflection of an order $n > 0$) is relatively constant for varying source distance *[Bla83]*.

### Spectral shape

Under certain circumstances, sound source spectrum varies as a function of distance. At greater distance (above $15\ m$ *[Bla83]*), the sound absorbing properties of air significantly modify the higher frequencies of the source. Moreover, these properties depend on environmental factors such as relative humidity or the temperature. *[Ing53]* suggested that at $40\ \%$ of humidity, the attenuation peak was at $4000$~Hz and was about $6\ dB$ every $100\ m$. Some studies suggested that humans take advantage of binaural cues in their distance judgement. *[Col68]* showed that perceived distance varies when you cut off the high frequencies of an click stimulus. He tested several distances (from $2.5$ to $8.5\ m$) and observed that for closer source the perceived distance increases when you remove high frequencies (above $7680\ Hz$). For further sources, the perceived distance is roughly accurate. But these results are challenged by several other studies such as *[Koe00][CTS68]*

---

**Todo**

These last two articles need to be read more deeply.

---

### Other factors in distance perception

**Vision** is known to affect percept of auditory space, including perceived distance.

**Familiarity** and prior information about the characteristics of a sound can significantly influence the auditory distance perception.

## Dynamic cues

As we briefly explain above, localisation can be improved or remove disambiguation through head movements and hence dynamic cues changes either by a source movement or a listener's movement.

For localisation of sound source in space, a listener naturally seeks to orientate his head toward this one and face it. It is in that position that sounds are localise the most accurately. However, *[PN97a]* suggested that an improvement of localisation accuracy in azimuth can be obtained by dynamic cues even if the sound is too short for the listener to face it. This result showed that localisation cues called "dynamic" introduced by head movements contribute in themselves to the localisation percept of a source. According to *[Mac09]*, head movements from $5°$ (at $50°/s$) generate usable dynamic cues. This is why head movements are beneficial even for short sound as described by *[PN97a]* comparing a localization performances of a low-pass noise stimulus lasting $3$ or $0.5$ seconds with or without slight head movements. The front/back ambiguities are reduced by analysing the dynamic changes of *ITD* and *ILD*. For example, for a source in front of the listener. If the listener turn his head to the right along the horizontal plan, the sound source will be perceived closer to the left ear. If he turn his head to the left, the sound source will be perceived closer to the right ear. If the source is behind the listener's head, the effect will be the opposite

---

**Todo**

create a figure explaining that.

---

*[PN97b]* studied the effect of dynamic cues in the elevation plan and suggested that head movement in this plan are beneficial for sources really high or low (±30°). *[Wal39][Wal40]* explained this by the fact that in these conditions the amplitude of dynamic variations of interaural cues lead by the head rotations are lower than sources closer of the horizontal plane. By using a low-pass noise, *[PN97b]* suggested that *ITD* changes are more reliable than *ILD*.

# CHAPTER 2

# The Filehne experiment

## Motivations

Speed perception has been intensively studied in vision. Even if the behaviour of speed mechanisms is still on debate (*[Fre01][FCW10]*), it exists low motion mechanisms that can extract the speed information.

**Todo**

mention the difference between speed and velocity. Meaning velocity contains speed and direction.

In audition, speed seems to be a difficult cue to extract and several findings suggested that audition doesn't have low level mechanisms but can still extract the information *[Gra86][MG91]*. We want to understand how speed perception is affected in audition when head movements occur and compare the results with vision findings. A famous illusion named after his author *[Fil22]* showed how speed perception is affected in vision when eyes movements occur.

During eye movements, the world around us remains perceptually stable despite of the retinal image slip (see Fig. 2.1). The pursuit adds motion to the image, hence, the brain must add this new estimate to the image motion in order to recover the object motion. This process doesn't work accuratly resulting in misperception of the object velocity during pursuit. This has been shown through several illusions such as the Aubert-Fleishl phenomenon (*[Aub86]*) where the pursued stimulus appears slower or the Filehne illusion (*[Fil22]*) showing that stationary objects appear to move. We will discuss the latter below and its impact on audition.

This illusion was named after the research who found it (*[Fil22]*). The illusion showed that a stationary object appears to move against the eyes movement. This process imply two estimates:

- the retinal image motion,
- the occulomotor system feedback known as *ERS*.

When we make a smooth eye movement to track a moving object, the visual system estimates the eyes velocity (using the *ERS*) and then substract it from the observed retinal motion Fig. 2.2.

As shown on the retinal image motion and the eye muscles feedback goes in opposite direction during the smooth

Fig. 2.1: Motion perception with or without eye pursuit of a moving object. The first image shows the perceived motion during eye fixation. The second shows the percieved motion during an eye pursuit.

Fig. 2.2: Signals used to infer the motion of an object during an eye pursuit.

pursuit. In order to obtain the object as stationary, these two estimates as to be equal.

$$\widehat{H} = \widehat{R} + \widehat{P}$$

| vision | Audition |
|---|---|
| Eyes rotation | Head rotation |
| Dot | Noise |
| Grating background | No background |
| No visual reference | No auditory **and** visual reference |

Fig. 2.3: Filehne illusion. Estimation of the speed of an object $\widehat{H}$ through the estimates of the eye pursuit $\widehat{P}$ and the retinal image motion $\widehat{R}$.

**Todo**

Find a way to insert a caption for this table. The caption should be the following: Equivalences between visual and auditory Filehne experiement.

**Todo**

Equivalence have no reference in the text at the moment, need to be fixed

## Broadcasting and motion of the acoustic signals

$$G = \sqrt{\exp\left(-2 \times \frac{x-p}{w}\right)^2} \tag{2.1}$$

In order to create a smooth motion we decided to have one signal per speaker and apply a spatial gaussian window letting us to compute the gains to apply on each channel for a given source position. In order to avoid phase problems at the listener's head, we used on each channel independent random gaussian noises. The spatial window is computed with a gaussian function (shown on (2.1)).

The gain for each channel is given by $x$ the position in degrees of each speaker, $p$ the position of the source and $w$ the width (spread or standard deviation) of the source in degrees. If $w = 0$, the source will be very ponctual[1], if $w > 0$, will be broadcast on several speakers. The position of the source is discrete with a $0.1°$ step. This is enough to obtain a perceived smooth and homogeneous movement and is much lower than the best *MAA* of $1°$ in front of the listener (*[Mil58]*) and consequently of the *MAMA* that is around $1°$ or larger (*[SP90][CG92][SMP92]*). One limitation of this technique is related to the physical distance between the speakers and corresponds to the parameter $w$ of the equation (2.1). The parameter $w$ can't be lower than the minimum distance between two loudspeakers. In this particular case, the motion will not be smooth anymore but will jump from one speaker to another. Another limitation is the computer's processor. Because the experiment has a real time constraint (due to the acquisition of head position data), the filtering process can disrupt the real processing.

[1] By ponctual, the source will be broadcast by the closest speaker and all over will be set at 0 dB.

Fig. 2.4: Spatialisation of the stimuli using an array of loudspeakers. Intensity of each speaker is respect to the gain of a gaussian function. These gains change over time.

## Head motion, the pursuit

A key point of the experiment is to control the participant head movement in order to keep his head speed as constant as possible. In vision, we know that eye movements are saccadic they can move smoothly when pursuit. First we tested on ourselves our capaticy to move our head at constant speed. It appeared that it was a very difficult task. It has been decided to lead a small and informal experiment in order to find the best method to obtain smooth head movements. We measured 6 participants using a metronome. The metronome used a click stimulus.

**Todo**

Nature and description of the conditions.

Participants were asked to anticipate the stimulus by pointing their nose at the click locations. A trial corresponded to two back and forth of the head The results showed mainly saccadic behaviors not related to the speed condition and not constant over time.

**Todo**

number of trial per session.

We decided then to use a pursuit noise that participants have to follow by pointing their nose at it. In order to help them to differenciate the test itself from the pursuit, a low-pass filter were applied on the pursuit.

## Auditory Filehne experiment

### Paradigm

The aim of the experiment was to examined the auditory motion perception during head movements. The general task took the form of a *2AFC* in which the subject was required to indicate which direction the stimulus appeared to move. Each trial was decomposed in two parts:

- the pursuit,
- the test.

Each subject participated to 4 sessions containing each three blocks. Before the first session[2], a training was carried out to familiarise participants with the task. Each session corresponded to three head speed conditions: 20, 40 and 60 $^\circ/s$. One block contained 140 trials and lasted about 30 minutes. Hence, one participant performed 1680 trials over 6 hours of experiment. Participant were free to choose how many blocks they want to do each time. If they chose to do at least two blocks, a rest of 5 mins were given between each block.

**Todo**

Why we decided to use this type of pursuit and another one ? Because the equivalent of a moving dot is a moving sound but with the problem of a non finite width, we choose to use a low pass filter to limit the interferences with the

[2] The participant, if necessary could ask for a training for following sessions because sessions occurs overs two weeks.

test and the we were obliged to stop the pursuit in order to not interfere with the test. In vision, usually use judge the background and not the the dot.

---

The pursuit in each condition lasted 3 seconds. The information about the pursuit are shown on Table 2.4.1 and Fig. 2.6. In order to balance the experiment, the pursuit direction was alternated on each trial.

The participant had to follow the pursuit by pointing is nose at it. This lasted two sweeps (back and forth), then the subject had to make a third sweep by himself. During this time, the test was presented and the participant had to judge his direction. The test was presented in order that both the test and the participants head should cross the 0° at the same time (Fig. 2.5).

| Condition ($°/s$) | Duration ($s$) | Displacement range (°) | Total displacement (°) |
|---|---|---|---|
| 20 | 3 | $\pm15$ | 60 |
| 40 | 3 | $\pm30$ | 120 |
| 60 | 3 | $\pm60$ | 180 |

Fig. 2.5: Process of the experiment over time and angular position of the head. The black plain line represents the head movement when the pursuit stimulus is on. The Black dashed line, the head movement when the pursuit is off. The blue thick line represents the test presentation.

---

**Todo**

caption to put with the table Head pursuit information regarding each condition such as total duration, displacement range (one head sweep) and total displacement.

---

The test was randomised on each trial using a range of duration from 400 to 600 $ms$. A range of 5 speeds with a step of $8°/s$. The basic range was from $-24$ to 24 $°/s$. After a preliminary analysis of the training, it was decided to shift the range of speeds in order to get a *PSE*. In order to prevent participants to make judgements according to the start and end of the stimulus (*[CB02]*), the test has been roved and its center varied between $\pm7.5$ ° (as shown on Fig. 2.6).

Fig. 2.6: Description of the experiment in terms of source and head displacement. The head movement according to the conditions will have maximum displacement of 90° centred on 0° (@ $60°/s$). The source will displacement is changing randomly from trial to trial and it's centre is always between $\pm$ 7.5°

---

**Todo**

- Talk about the intensity experiment that did not work until now
- change the different inkscape figure by their tikz equivalent
- save in a different folder, all script generating tikz plot from octave in a specific folder

---

## Analysis

On the six subjects, everyone completed the task required. Nevertheless, the analysis revealed that two of these participants had a strange behaviour and showed the biggest effect regarding the other participants. Outliers were defined as no head motion during the test stimulus presentation and as data acquisition problem. Per session, on average, there is about 0.13% of outliers with a maximum of 3 outliers on a session and a minimum of 0. This low percentage of trial rejection is explained by the observation of head movements on average and decided to keep almost all trials to lead an analysis based on true head movements. Results have been computer on each session and then averaged to get PSEs.

## Head movements

Head movements were driven by an audio pursuit target, but like eye movements to a lesser extent, they tends to be saccadic even when pursuit. To reduce this effect, a Savitzky-Golay filter (*[SG64]*) was applied on each trial. This process is achieved by using a local least-squares polynomial approximation (approximation of the second order in our case) resulting to a low pass filter on the data set[3].

Fig. 2.7 shows a typical head movement on a trial. The ideal head movement describes a triangle signal in order to keep a constant speed over time and angular displacement. Nevertheless, participants showed difficulties to reproduce correctly this pattern. This is explained by several reasons. Firstly, a typical participant pattern is a sinusoidal signal. The change of head direction can't be immediate due to the weight and inertia of the head. This effect add a delay to the pursuit. The other problem is poor width definition of an audio source. This prevents a good pursuit of the source. Because participant were in the dark with no visual cue, they can't use speakers or other references to stop or anticipate direction changes. This explains why the angular displacement of participant's head can be lower or greater than the ideal pattern and add another delay. Nevertheless, as shown by the figure Fig. 2.7, during the phase between head direction changes, the participant is able to keep his head movement quite steady.

**Todo**

it could be interesting to compute the percentage around the speed target

Fig. 2.7: Head tracking during a trial at condition 20 $°/s$. orange plain line represents the ideal head movement over time and angular displacement. The blue plain line represents the head movement of participant 1 during the trial 4 of session 1.

In order to extract only smooth pursuit movement during both sweeps of the pursuit task. It has been decided to keep only 1 second of signal when the head is centered on $0°$ (see Fig. 2.8). Then, for each condition and participant, the mean speed has been computed on each trial and then averaged across all sessions for the pursuit and test. The results are shown on Fig. 2.9.

Fig. 2.8: Head pursuit speed computation. The grey zones represent the meaningful parts of head movements used to compute the head speed during pursuit.

The difficulty of participant to follow the pursuit is confirmed by the left hand side figure that shows the average for each participant and condition during the pursuit presentation. At $20°/s$ participant are relatively close to the target wheras for $40$ and $60°/s$ the general behaviour is to slow down the head speed regarding the target. Nevertheless, participants 4 and 6 tends to keeps their head around the same speed whatever the target is and both are around $50°/s$. Even if they understood the task, these participant seems to have difficulties to extract the speed information of a moving source and can't use or make the difference between several sets of interaul cues. If a subject follow perfectly a sound source, the pair of *ITD* and *ILD* will not evolve over time[4]. Based on these cues, a subject should be able to tell if he is late or ahead regarding the sound source. These cues are the only cues available during this task and participant 4 and 6 seems to not be able to use in a accurate way these cues.

**Todo**

[3] For a better understanding of this type of filter, the reader can refer to *[Sch11]*.

[4] Or at least in a insifignant way, with small reflections due to the torso.

These pursuit information are not accurate enough because of the extraction method used. I need to correct that in order two possible ways: either try to find the 0 deg and extract 1 second of signal around it or transform the signal in order to keep all the meaningful information.

---

On the right hand side figure is shown average speed for each participant and condition. The global behaviour is that all participant accelerate their head movements. This suggests that, even with a reference before each trial, subject can't keep the same head speed. The change can be up to $25°/s$, that is a radical change between two head sweeps.

Fig. 2.9: Head speed distribution according to participants and speed conditions. The left figure represents mean head speeds during the pursuit and right one represents the mean head speeds during the test presentation. For the pursuit, only sweeps without head direction changes was kept.

## perceived speed

What is the impact of the head movement on the perceived speed of the test. As a reminder, participant were asked to judge the direction of the test presented while they were moving their head. The only criteria modified during the task was the speed of the test. And this task was led for 3 head speeds conditions. To analyse the data, for each session, participant and condition, the percentage of test perceived in the direction of the head was computed. Then a psychometric function was extracted using a Probit analysis (*[Fin71]*. The meaningful information is the *PSE* at $50\%$ representing the perceived stationnarity of the test. The figure Fig. 2.10 shows the results of participant 1 for his first session on each condition. We can observed firstly that all three *PSE* are above the $0°/s$. If someone makes a head movement in front of a fixed sound source, if no effect, were perceived, the perceived speed of the sound source should be $0°/s$. In the present case, there is a compensation from the participant and the compensation is in the opposite direction to the head. This corresponds to a Filehne illusion as described by *[Fil22]*. This suggests that participant 1 makes an estimation error that would maybe be on the proprioceptive information ($\widehat{H}$) or in the cochlear image motion information ($\widehat{I}$) as suggested in vision by *[FB98]*. Secondly, the figure suggests that the Filehne illusion increased with the head speed according to each condition.

---

**Todo**

Comment: Nevertheless, as shown on Fig. 2.9, participant does not necessary match the theoric head speed conditions espacially during the test presentation. In order to confirm the effect, the Fig. 2.11 shows the *PSE* of each participant for each condition. But instead of plotting the theoric head speeds, it's the actual head speeds that are shown. All participant, whatever the their head speed is suffer the illusion in the same direction (opposite to the head movement). Moreover, the illusion increases as the head speed increases for all participant. An interesting observation would be that the illusion tends to evolve linearly with respect to the head speed. This is difficult to verify as the number of participant is really low. Indeed participant 2 and especially participant 4 doesn't show a linear illusion but it could explained by the fact that their behaviour were a bit strange compare to the othertodo{really badly explained, need to be rewritten with a better explaination (maybe show their psychometric functions for left and right). As shown on Fig. 2.10, the psychometric function means that if the participant makes a head movement across a static auditory object, this object would appear to move in the opposite direction of the head movement.

---

Fig. 2.10: Psychometric function of the participant 1 for one session. The psychometric function shows the PSE of the test velocity according the test stimulus perceived in the direction of the head. At the 50 $\%$, the stimulus appeared to be stationnary. each color represents one condition (20, 40 and 60 $°/s$).

Fig. 2.11: Individual differences of PSEs according to the actual speeds on each condition for each participant.

**Todo**

need to talk a bit about the shift problem and the anova ran on Cass data.

# Discussion

## Filehne experiment improvements

What could be improve for this experiment? Get more participants in order to ensure the results and especially confirm the results given by participant 2 and 4. Improve the pursuit system by putting the speaker closer and reduce the theoritical size of the source, and maybe use a higher bandwidth in order to get a more ponctual sound source easier to follow. Another idea would be to trained a lot people to excute head movement at specified speed by given them an auditory feedback if they are too slow or fast.

## Stimulus properties

According to the results observed above, all subjects suffered the same effect at different strenghs and whatever their actual head speed was. In other words, the audio Filehne effect means if someone moves his head in front of a fixed sound source, this latter will appear to move in the opposite direction of the head movement. Moreover, this effect seems dependent of the head speed. Based on these result, we can assume that audition speed perception, like in vision, will be affected function of several properties. According to *[FB98]*, the retinal image will be affected by several stimulus properties. In the auditory domain, the cochlear image ($\widehat{I}$) will be affected by stimulus properties ($\Omega$) as given by the (2.2).

$$\widehat{Object} = \widehat{I}(\Omega) + \widehat{H} \tag{2.2}$$

If one of the two estimates evolves in one direction of the other, then the perceived velocity will change[5]. By changing a propertie of a stimulus, the cochlear image estimate ($\widehat{I}$) would change (see Table 2.6.2 for equivalence). *[HCTM07][VPC08]* suggested that a visual pattern will appear to move faster at lower intensities. Hence, the Filehne illusion increases as the luminance decreases. Luminance is a visual properties that as an equivalent in audition called the intensity.

| Vision | Audition |
|---|---|
| pursuit ($\widehat{P}$) | neck, vestibulus ($\widehat{H}$) |
| retinal image motion ($\widehat{E}$) | cochlear image motion ($\widehat{I}$) |

Fig. 2.12: Perceived speed function to the intensity. The blue line represents the visual tendency according of results given by *[HCTM07][VPC08]*. The purple plain line represents the possible behavior in audition that would be the opposite of vision after informal tests.

Filehne vision versus audition estimates.

**Todo**

check the letters for each estimates and find a way to put a caption under the table.

Following the above results, if the intensity of a sound increases, the resulting speed perception would decrease. Nevertheless, informal tests on the author and colleagues suggested that the perceived speed should increase with

[5] In this particular case, the illusion could increase, decrease or be invert (as suggested by *[FB98]*, that's why velocity is used instead of speed.

respect of the intensity. Unfortunately, pilot data showed on two naive participants reported that participant were unable to do the task. It seems that they were unable to make judgment and can't linked perceived speed and intensity.

---

**Todo**

need to be a bit more precise on the paradigm and why participant were unable to do the task... difficulties to use the speed as the only cue, need to check the exact paradigm.

---

As explained in the section *Auditory localisation*, localisation cues are really important and are function of the frequency (Fig. 2.6.3). Localisation is usually improved when all localisation cues are available. This means that speed perception could be affected is the spectral components of the source contains only one or two localisation cues in it. *[Mil58]* showed that *MAA* is more affected in a specific spectral zone where *ITD* and *ILD* aren't effective enough and where the *MAA* in front of the listener increases up to 3° (Fig. 2.13).

This result suggests the possibility that the perceived speed would increase more if a source with a bandwidth from 2 to 3 kHz is presented where human can rely on *ITD* or *ILD*.

## Correlation between audio and visual Filehne illusion

In order to estimate the object motion during pursuit in vision is to combine estimates of eye velocity and retinal motion and in audition of head velocity and cochlear motion. If in both cases, the combination of the estimates happened in an early stage, auditory and visual Filehne illusion should be independent. Nevertheless, some recent works such as *[KPB+03]* suggested that 'retinal' and 'extra-retinal' motion pathways shared a common noise source suggesting that observers do not have a direct access to the retinal motion and that the combination of the estimates should happenned in a later stage of the perceptual system. This have been confirmed by *[FCSS09]* who used a *2IFC* task in which observers had to indicate which interval contained the faster background motion, while pursuing a target that moved across the background.

Hence, it would be interesting to lead in parallel both auditory and visual Filehne illusion experiments and observe if a correlation between both data sets exists (Fig. 2.14). If so, it will suggest that both auditory and visual motion pathways are shared in a later stage of the perceptual system and confirm results given above.

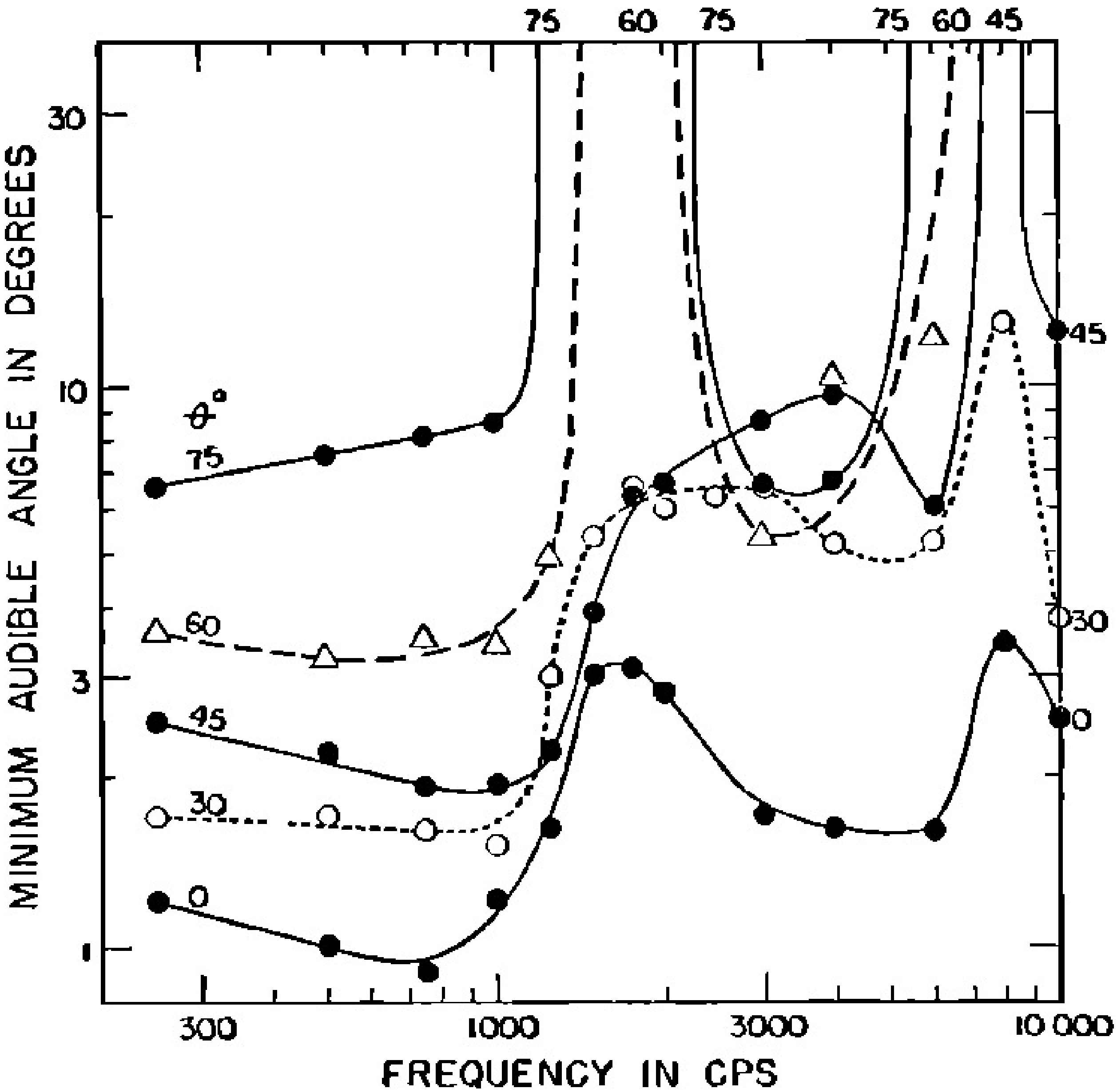


Fig. 2.13: Frequency dependence of localisation blur in azimuth (expressed here as “Minimum Audible Angle”) using pure tones, as a function of the sound source azimuth position $\theta$. (After [Mil58]).

Fig. 2.14: Possible correlation between the visual and auditory Filehne illusion.

CHAPTER 3

# Laboratory

In order to measure the phenomena that we were interested in (see chapter *The Filehne experiment*). We created for the Perception group of the School of Psychology of Cardiff University a new audiovisual lab. In audio research, there is no standard measurement system but according to the needs we will give a priority to two main techniques:

- *VAS*
- *RAS*

## Virtual Auditory Space vs Real Auditory Space

The *VAS* is the ability to create the illusion of any free-field environment using a closed-field sound system such as headphones or loudspeakers. This technique assumes that identical stimuli will be perceived identically at a listener's eardrum whatever the physical mode of delivery. It is now accepted that the simulation of acoustical space is best achieved using closed-field systems since headphones allow a complete control over the signal delilvered to the listener's eardrums. The disadvantage of this technique is that it requires compensation of the transfer function of the sound delivery system itself. Moreover, in order to give to the listener the perfect illusion of a 3D audio scene, you will need to use the binaural technique. To achieve that, it is necessary to recreate at each ear, the signals that would be perceived naturally. The use of the *HRTF* is the best way to reproduce the localisation cues needed.

### Binaural broadcasting technique

The binaural synthesis is based on the use of the pair of binaural filters obtained from the *HRTF*. At each source position in the space $r, \theta, \phi$ it exists a pair of *HRTF*, that we can obtain through a model or a set of measurements. In order to place a virtual source at a given position, it is necessary to find the pair of *HRIR* corresponding to the position in a database if available or calculate the interpolation and deduce a pair of binaural filters $x_L$ and $x_R$ adapted to the chosen implementation. For the headphone diffusion, the simplest way is to convolve the monophonic and anechoic signal $x$ with each filter in order to obtain the signals $x_L$ and $x_R$ that will be broadcast on the headphones (see Fig. 3.1). In addition, it is necessary to compensate for the headphone that act as a filter.

The spectral filtering of a sound source before it reaches the eardrum is called the *HRTF*. The binaural *HRTF* can be thought of as a frequency-dependent and amplitude and time-delay differences that result primarily from the complex

Fig. 3.1: Binaural techique on headphones. After *[Gui09]*.

shaping of the pinnae. *[Bat67]* claimed that the folds of the pinnae cause time delays within a range of $0$ to $300\mu$. This is a cause of a significant change in the spectral content at the eardrum. Because of the asymetric shape of the pinnae, this spectral changes vary with the source position. Moreover, the shape of pinnae differ from one subject to another. This means that in theory, we should measure the *HRIR* for an infinite number of positions in order to reconstruct perfectly the signal at the eardrums. Because it is impossible to measure an infinite number of points and because, measuring impulse responses of a subject is still nowaday is difficult and long task suggesting a sampling of a finite number position and then interpolate the missing positions. Another way is to use a bank a average *HRTF* and use the same bank for all subject. Both techniques bring artefacts once convolved with the signals. Results are localisation and externalisation of sounds problems. The externalisation problem is not still perfectly known. Nevertheless, *[Gui09]* suggested several possibilities that could have an impact on the externalisation such as the fact that the listener knows that signal is broadcast through the headphones, and feel the pressure of it on his ears. The absence of visual cues, or incoherent signals between the visual and audio modalities. The acoustics signals at the eardrums can be as well degrade due the the distortion brought by the headphones.

### Multi loudspeakers technique

The use of loudspeakers instead of headphones avoid troubles about externalisation of the sound and a difficult *HRTF* measuring process. Spatialisation of sound is more robust, all spatialisation cues are naturally available and don't need to be recreated. Nevertheless, several problems still exist such as the interpolation of sounds located between two speakers.

**Todo**

Becareful, in both cases (VAS and RAS), the interpolation is not a real problem for the simple reason that in VAS, we can't measure an infinite number of points, hence, we will interpolate several positions. In RAS, we will not have an infinite number of speakers, thus, we will interpolate any position that is located between two speakers.

### Multi loudspeakers technique

The use of loudspeakers instead of headphones avoid troubles about externalisation of the sound and a difficult *HRTF* measuring process. Spatialisation of sound is more robust, all spatialisation cues are naturally available and don't need to be recreated. Nevertheless, several problems still exist such as the interpolation of sounds located between two speakers.

**Todo**

Becareful, in both cases (VAS and RAS), the interpolation is not a real problem for the simple reason that in VAS, we can't measure an infinite number of points, hence, we will interpolate several positions. In RAS, we will not have an infinite number of speakers, thus, we will interpolate any position that is located between two speakers.

## Equipment

Visual motion has been intensively investigated and need a quite standardised equipment (see *[KB10][BJVDB01][Fre01]*). Audio motion requires *ad hoc* systems and can differ a lot from one lab to another and will depend mainly on using *VAS* or *RAS* (*Virtual Auditory Space vs Real Auditory Space*) and many other parameters. The lab's wiring diagram is given on Fig. 3.2 and a picture of the result is given on Fig. 3.3.

Fig. 3.2: Schematic of the lab audiovisual system. In green are represented the inputs, in brown the outputs.

Fig. 3.3: Photo of the laboratory with a dummy head instead of a participant.

## The room

is a parallelipedic shape with a superficy of $13.76m^2$ ($3.2 \times 4.3m$). The lab has several characteristics such as:

- black walls in order to minimize light reflections,
- a proof-sound material on the wall to minimize acoustics reflections,
- no isolation from the outside noise.

A plastic rail surrounding the room at the ears heigh (when a participant is seated) has been covered with foam in order to reduce its impact on the acoustic. A measure of the *RT* gave a result of $60ms$ on average. A measure of the noise floor has been done and gave a result of 30 dB on average with a pic around 60 dB at 200 Hz corresponding to the cooler system when it is turned on (see Fig. 3.4). Further investigation using acoustic antenna technique (such as beamforming or holography) would help to find where is the noise position and correct it in order to lower that noise. Because it is quite low frequency, it should not be perceived as a ponctual source by the participants and not interfere in the experiments.

Fig. 3.4: Noise floor of the laboratory with cooler system on.

## Loudspeakers

For the broadcasting of the signal, we needed multiple loudspeakers using a *RAS* (see *Virtual Auditory Space vs Real Auditory Space*). Given the constraints we decided to use broadband speakers with a small size in order to have a quite high density. The system is composed of 24 *Minx min 10*, *Cambridge Audio* loudspeakers (see *[Cambridge Audio11]*). These speakers are passive and measure $80 \times 80 \times 80mm$. The system uses 22 fixed speakers (with 2 speakers that can be placed where it is needed) along an hemicircle with a distance between each speaker of $7.5°$. As shown on the Fig. 3.5, the bandwith of the speakers is on average about from 200 Hz to 10000 Hz. This is enough to use white noise in order to be able to use all acoustic available cues.

Fig. 3.5: Frequency response of the speaker 12 (placed @ $0°$).

## Amplifiers

Because we decided to keep amplifiers in the room, we needed a passive cooling system. We chose four 6 channels *AMP-CH06, Auna* amplifiers:

- Electric power: 570 Watts RMS,
- frequency response: 20 to 20000 Hz,
- *SNR*: 95 dB,
- impedance: 16 Ω.

## Head tracking

In order to measure head tracking, we have two systems that is used according to the constraints of the experiment. A magnetic head tracker *Flock of Birds, Ascension* (see *[Ascension04]*) is used to record accurate head movements position and rotation in 3 dimensions. This tracker let us to record information in real time if it is needed to change the behavior of the experiment according to the head movements. If the participant can't be aware of his head tracking, a webcam *LifeCam HD 3000, Microsoft* (see *[Microsoft11]*) fixed above the participant's head on the ceiling is used to record and movement and is analysed afterwards. This system is less accurate and record only rotation in one dimension and position in 2 dimensions.

## Video projector

In order to lead multi modalities experiments such as audiovisual experiments, a video project has been installed. Because of the room characteristics, a small and quiet projector were needed. A *Qumi Q2, Vivitek* (see *[Vivitek13]*) has been chosen and will be fixed on the ceiling above the participant's head.

## Sound card

for flexibility we used a *24~I/O, Motu* *DAC* and a *PCIexpress, Motu* sound card (see *[Motu13]*). The sound card can handle up to 4 *DAC* (96 channels) at 24 bits quantification and 96 kHz.

## IT equipment

The computer is in a operating room next to the lab in order to minimise the acoustic impact. The main components of the computer are a *i5-2400, Intel* processor with 3 GB of *RAM*.

## Softwares

Any software capable of using *ASIO* driver can be used to handle the high number of channels if there is no need of head tracking. Nevertheless, for the processing and for the experiments described in this document, *Pure Data* has been used to lead the experiments, *Matlab, Mathworks* or *GNU Octave* with the toolbox *Playrec* has been used for measurements or data analysis. The main advantage of using *Pure Data* is the real time processing and its capacities to handle the head tracker *Flock of Birds, Ascension*.

## Loudspeaker compensation

As shown on Fig. 3.5, the response of the speaker is chaotic and because of its mechanic assembly, the frequency response will differ from one to the other. These differences can be heard by the participants and give them intrusive spectral or intensity cues that could bias the experiments. Because of the spectral response of the speakers, rather than trying to flatter it, it has been decided to bring the same default to every speakers. The speaker at $0°$ in front of the listener is the reference. The principle is to extract for each speaker impulse response the corresponding excitation pattern[1] (see equation (3.1)), get the spectrum difference from the reference excitation according to the current one and convolve the current impulse response with the spectrum difference.

$$W(g) = (1 + pg)\exp(-pg) \tag{3.1}$$

Where $p$ determines the shape of the pass band filter. $g$ is the deviation in frequency from the filter center frequency divided by the center frequency.

[1] The excitation pattern is the distribution of internal excitation as a function of some internal variable related to frequency.

CHAPTER 4

# Bibliography

CHAPTER 5

# Glossary

**2AFC** Two-Alternative Forced Choice.

**2IFC** Two-Interval Forced Choice.

**AAM** Auditory Apparent Motion.

**AMAE** Auditory Motion AfterEffect.

**ASIO** Audio Stream Input/Output.

**ASW** Apparent Source Width.

**DAC** Digital Analog Converter.

**DS** Direction Specific.

**ECS** Extra Cochlear Signal.

**ERS** Extra Retinal Signal.

**HOA** Higher Order Ambisonics.

**HRIR** Head Related Impulse Response.

**HRTF** Head Related Transfer Function.

**IC** Interaural Coherence.

**ICC** Inter-Channel Coherence.

**ILD** Interaural Level Difference.

**IPD** Interaural Phase Difference.

**ITD** Interaural Time Difference.

**JND** Just Noticeable Difference.

**MAA** Minimum Audible Angle.

**MAE** Motion AfterEffect.

**MAMA** Minimum Audible Movement Angle.

**PSE** Point of Subjective Equality.

**RA** Research Assistant.

**RAM** Random Access Memory.

**RAS** Real Auditory Space.

**RT** Reverberation Time.

**SD** Standard Deviation.

**SNR** Signal to Noise Ratio.

**VAS** Virtual Auditory Space.

**VBAP** Vector Based Amplitude Panning.

**WFS** Wave Field Synthesis.

# Bibliography


[Ascension04] *The Flock of Birds*. Ascension, 2004. Installation and operating guide. URL: ftp://ftp.ascension-tech.com/MANUALS/Flock_of_Birds_Manual-RevC.pdf.

[Microsoft11] *LifeCam HD-3000*. Microsoft, 2011. Technical data sheet.

[Cambridge Audio11] *Minx Min 10 loudspeaker*. Cambridge Audio, 2011. Technical specifications. URL: http://www.cambridgeaudio.com.

[Motu13] *Motu PCI-424*. Motu, 2013. Manual.

[Vivitek13] *Qumi Q2*. Vivitek, 2013. Data sheet.

[Aub86] Hermann Aubert. Die bewegungsempfindung. *Archiv für die gesamte Physiologie des Menschen und der Tiere*, 39(1):347–370, 1886. doi:10.1007/BF01612166.

[Bat67] *The role of the pinna in human localization*, volume 168 of B, Biological Sciences, The Royal Society of London, 1967. doi:10.1098/rspb.1967.0058.

[Bla83] Jens Blauert. *Spatial Hearing: the psychophysics of human sound localization*. MIT Press, second (revised in 1997) edition, november 1983. URL: http://books.google.co.uk/books?id=wBiEKPhw7r0C&lpg=PR1&pg=PR1#v=onepage&q&f=false.

[BJVDB01] Eli Brenner, eroen B. J. JSmeets, and A. V. Van Den Berg. Smooth eye movements and spatial localisation. *Vision Res.*, 41(17):2253–2259, 2001. doi:10.1016/S0042-6989(01)00018-9.

[Bur58] J.F. Burger. Front-back discrimination of the hearing system. *Acustica*, 8:301–302, 1958.

[CB02] S. Carlile and C. Best. Discrimination of sound velocity in human listeners. *J. Acoust. Soc. Am.*, 111(2):1026–1035, 2002. doi:10.1121/1.1436067.

[CG92] David W. Chandler and D. Wesley Grantham. Minimum audible movement angle in the horizontal plane as a function of stimulus frequency and bandwidth, source azimuth, and velocity. *J. Acoust. Soc. Am.*, 91(3):1624–1636, 1992. doi:10.1121/1.402443.

[CTS68] Paul Cochran, Janet Throop, and W. E. Simpson. Estimation of distance of a source of sound. *The American Journal of Psychology*, 81(2):198–206, 1968. URL: http://www.jstor.org/stable/1420327.

[Col68] Paul D. Coleman. Dual role of frequency spectrum in determination of auditory distance. *Journal of the Acoustical Society of America*, 44(2):631–634, 1968. doi:10.1121/1.1911132.

[Dan11] Adrien Daniel. *Spatial auditory blurring and applications to multichannel audio coding*. PhD thesis, Université Pierre et Marie Curie, 2011.

[Edw55] Austin S. Edwards. Accuracy of auditory depth perception. *Journal of General Psychology*, 52:327–329, 1955. doi:10.1080/00221309.1955.9920247.

[Fil22] W. Filehne. Uber das optische wahrnehmen von bewegungen. *Zeitschrift für Sinnephysiologie*, 53:134–145, 1922.

[Fin71] David J. Finney. *Probit analysis*. Cambridge University Press, 3rd edition edition, 1971.

[FF68] H. Geoffrey Fisher and Sanford J. Freedman. The role of the pinna in auditory localization. *Journal of Auditory Research*, 8(1):15–26, 1968.

[FB98] T. C. Freeman and Martin S. Banks. Perceived head-centric speed is affected by both extra-retinal and retinal errors. *Vision Research*, 38(7):941–945, 1998. doi:10.1016/S0042-6989(97)00395-7.

[Fre01] Thomas C. A. Freeman. Transducer models of head-centred motion perception. *Vision Res.*, 41(21):2741–2755, 2001. doi:10.1016/S0042-6989(01)00159-6.

[FCW10] Thomas C.A. Freeman, Rebecca A. Champion, and Paul A. Warren. A bayesian model of perceived head-centered velocity during smooth pursuit eye movement. *Current Biology*, 20(8):757–762, 2010. doi:10.1016/j.cub.2010.02.059.

[FCSS09] Tom C. A. Freeman, Rebecca A. Champion, Jane H. Sumnall, and Robert J. Snowden. Do we have direct access to retinal image motion during smooth pursuit eye movements? *Journal of Vision*, 9(1):1–11, 2009. doi:10.1167/9.1.33.

[Gra86] D. Wesley Grantham. Detection and discrimination of simulated motion of auditory targets in the horizontal plane. *J. Acoust. Soc. Am.*, 79(6):1939–1949, 1986. doi:10.1121/1.393201.

[Gui09] Pierre Guillon. *Individualisation des indices spectraux pour la synthese binaurale : recherche et exploitation des similarites inter-individuelles pour l'adaptation ou la reconstruction de HRTF*. PhD thesis, Université du Maine, 2009.

[HCTM07] Stephen T. Hammett, R. A. Champion, Peter G. Thompson, and Antony B. Morland. Perceptual distortions of speed at low luminance: evidence inconsistent with a bayesian account of speed encoding. *Vision Res.*, 47:564–568, 2007. doi:10.1016/j.visres.2006.08.013.

[Ing53] Uno Ingård. A review of the influence of meteorological conditions on sound propagation. *Journal of the Acoustical Society of America*, 25(3):405–411, 1953. doi:10.1121/1.1907055.

[Koe00] *Effect of distance on localization of complex stimuli*, The Association for Research in Otolaryngology, 2000.

[KPB+03] Anton E. Krukowski, Kathleen A. Pirog, Brent R. Beutter, Kevin R. Brooks, and Leland S. Stone. Human discrimination of visual direction of motion with and without smooth pursuit eye movements. *Journal of Vision*, 3(11):831–840, 2003. doi:10.1167/3.11.16.

[KB10] Kerstin Königs and Frank Bremmer. Localization of visual and auditory stimuli during smooth pursuit eye movements. *Journal of Vision*, 10(8):1–14, 2010. doi:10.1167/10.8.8.

[LB02] Erno H. A. Langendijk and Adelbert W. Bronkhorst. Contribution of spectral cues to human sound localization. *Journal of the Acoustical Society of America*, 112(4):1583–1596, 2002. doi:10.1121/1.1501901.

[Mac09] Erwan A. Macpherson. Stimulus continuity is not necessary for the salience of dynamic sound localization cues. *The Journal of the Acoustical Society of America*, 125(4):2691–2691, 2009. doi:10.1121/1.4784288.

[MG91] John C. Middlebrooks and David M. Green. Sound localization by human listeners. *Annu. Rev. Psychol.*, 42:135–159, 1991. doi:10.1146/annurev.psych.42.1.135.

[Mil58] A. W. Mills. On the minimum audible angle. *J. Acoust. Soc. Am.*, 30(4):237–246, 1958. doi:10.1121/1.1909553.

[PN97a] Stephen Perrett and William Noble. The contribution of head motion cues to localization of a low-pass noise. *Perception and Psychophysics*, 59(7):1018–1026, 1997. doi:10.3758/BF03205517.

[PN97b] Stephen Perrett and William Noble. The effect of head rotations on vertical plane sound localization. *J. Acoust. Soc. Am.*, 102(4):2325–2332, 1997. doi:10.1121/1.419642.

[Ray76] L. Rayleigh. Our perception of the direction of a source of sound. *Nature*, 14:32–33, 1876. doi:10.1038/014032a0.

[Ray07] Lord Rayleigh. On our perception of sound direction. *Philosophical Magazine*, 13(74):214–232, 1907. doi:10.1080/14786440709463595.

[Rum12] Francis Rumsey. *Spatial audio*. Taylor & Francis, 2012.

[SP90] Kourosh Saberi and David R. Perrott. Minimum audible movement angles as a function of sound source trajectory. *J. Acoust. Soc. Am.*, 88(6):2639–2644, 1990. doi:10.1121/1.399984.

[STFJ55] T. T. Sandel, D. C. Teas, W. E. Feddersen, and L. A. Jeffress. Localization of sound from single and paired sources. *Journal of the Acoustical Society of America*, 27(5):842–852, 1955. doi:10.1121/1.1908052.

[SG64] Abraham Savitzky and Marcel J. E. Golay. Smoothing and differenciation of data by simplified least squares procedures. *Analytical Chemistry*, 36(8):1627–1639, 1964.

[Sch11] Ronald W. Schafer. What is a savitzky-golay filter? *IEEE Signal Processing Magazine*, 28(4):111 – 117, 2011. doi:10.1109/MSP.2011.941097.

[SCSK00] Barbara G. Shinn-Cunningham, Scott Santarelli, and Norbert Kopco. Tori of confusion: binaural localization cues for sources within reach a listener. *Journal of the Acoustical Society of America*, 107(3):1627–1636, 2000. doi:10.1121/1.428447.

[SN36] S. S. Stevens and E. B. Newman. The localisation of actual sources of sound. *American Journal of Psychology*, 48(2):297–306, 1936. doi:10.2307/1415748.

[SMP92] T. Z. Strybel, C. L. Manligas, and D. R. Perrott. Minimum audible movement angle as a function of the azimuth and elevation of the source. *Human Factors*, 34:267–275, 1992. URL: http://apps.webofknowledge.com/full_record.do?product=WOS&search_mode=GeneralSearch&qid=4&SID=S15k8IbfkIBAkPmBc34&page=2&doc=17.

[Tho92] S. P. Thompson. On the function of the two ears in the perception of space. *Philosophical Magazine*, 13:320–334, 1892.

[TR67] Willard R. Thurlow and Philip S. Runge. Effect of induced head movements on localization of direction of sounds. *J. Acoust. Soc. Am.*, 42(2):480–488, 1967. doi:10.1121/1.1910604.

[VPC08] Maryam Vaziri-Pashkam and Patrick Cavanagh. Apparent speed increases at low luminance. *Journal of Vision*, 8(16):1–12, 2008. doi:10.1167/8.16.9.

[Wal39] Hans Wallach. On sound localization. *Journal of the Acoustical Society of America*, 10(4):270–274, 1939. doi:10.1121/1.1915985.

[Wal40] Hans Wallach. The role of head movements and vestibular and visual cues in sound localization. *Journal of Experimental Psychology*, 27(4):339–368, 1940. doi:10.1037/h0054629.

[WK99] Frederic L. Wightman and Doris J. Kistler. Resolution of front–back ambiguity in spatial hearing by listener and source movement. *Journal of the Acoustical Society of America*, 105(5):2841–2853, 1999. doi:10.1121/1.426899.

[WS54] Robert S. Woodworth and Harold Schlosberg. *Experimental psychology*. Holt, 3rd edition, 1954.

[You31] P.T. Young. The role of head movements in auditory localisation. *Journal of Experimental Psychology*, 14(2):95–124, 1931. doi:10.1037/h0075721.

# Index

## Symbols



## A



## D



## E



## H



## I



## J



## M



## P



## R



## S



## V



## W