Modulation frequency as a cue for auditory speed perception

Article in Proceedings of the Royal Society B: Biological Sciences · July 2017
DOI: 10.1098/rspb.2017.0673

CITATIONS
0

READS
73

3 authors:

Irene Senna
Ulm University
21 PUBLICATIONS 226 CITATIONS
See Profile

Cesare V Parise
Oculus VR, Greater Seattle Area
46 PUBLICATIONS 725 CITATIONS
See Profile

Marc O Ernst
Ulm University
226 PUBLICATIONS 6,766 CITATIONS
See Profile

Some of the authors of this publication are also working on these related projects:

Interactive Coaching in Virtual Reality View project

Human Multisensory Perception View project

All content following this page was uploaded by Irene Senna on 12 July 2017.
The user has requested enhancement of the downloaded file.
Modulation frequency as a cue for auditory speed perception

Irene Senna¹, Cesare V. Parise²,³ and Marc O. Ernst¹

¹Department of Applied Cognitive Psychology, Ulm University, Ulm, Germany
²Oculus Research, Redmond, WA, USA
³CITEC, Bielefeld University, Bielefeld, Germany

Electronic supplementary material is available online at rs.figshare.com.

Unlike vision, the mechanisms underlying auditory motion perception are poorly understood. Here we describe an auditory motion illusion revealing a novel cue to auditory speed perception: the temporal frequency of amplitude modulation (AM-frequency), typical for rattling sounds. Naturally, corrugated objects sliding across each other generate rattling sounds whose AM-frequency tend to directly correlate with speed. We found that AM-frequency modulates auditory speed perception in a highly systematic fashion: moving sounds with higher AM-frequency are perceived as moving faster than sounds with lower AM-frequency. Even more interestingly, sounds with higher AM-frequency also induce stronger motion aftereffects. This reveals the existence of specialized neural mechanisms for auditory motion perception, which are sensitive to AM-frequency. Thus, in spatial hearing, the brain successfully capitalizes on the AM-frequency of rattling sounds to estimate the speed of moving objects. This tightly parallels previous findings in motion vision, where spatio-temporal frequency of moving displays systematically affects both speed perception, and the magnitude of the motion aftereffects. Such an analogy with vision suggests that motion detection may rely on canonical computations, with similar neural mechanisms shared across the different modalities.

1. Introduction

In our daily lives, we often experience a relation between the speed of moving objects and the properties of the sounds produced by such motion. For instance, when manually exploring a corrugated surface, such as when sliding a hand on a fence, the contact with that surface generates rattling sounds whose frequency of amplitude modulation (AM) seems to correlate with the speed of motion, while the carrier frequency stays roughly constant: the faster the motion, the higher the AM-frequency of the rattling sound. Thus, in principle, the temporal frequency of a sound’s envelope might be used as a relative cue to estimate the speed of a sound source.

Over the last years, several studies have demonstrated how knowledge of statistical regularities in natural scenes can help perceptual systems reducing the complexity of sensory information and inferring the actual state of the world [1–6]. Through repeated exposure to statistical correlations across two or more sensory cues, the perceptual system learns such associations and builds up expectations about the natural mapping across the various cues [7–10]. The AM-frequency of a sound might be used as a relative cue to speed: indeed, assuming a constant environment (e.g. isotropic texture of a stimulus that does not vary over time), differences in the temporal rates of the sound produced by the contact between a moving object and a surface might be interpreted as differences in speed.

Besides audition, temporal frequency is a critical cue also in visual motion perception [11–12], and more recently it has been shown to also affect tactile speed perception [13–15]. For instance, the spatio-temporal frequency of a
grating systematically modulates the stimulus’ perceived speed. That is, perceived speed increases with increasing spatio-temporal frequency [11]. Moreover, temporal frequency affects the magnitude of the aftereffect induced by moving visual patterns [12]. In a series of psychophysical experiments, here we used moving rattling sounds to investigate whether AM-frequency induces analogous effects in audition. With this we would identify AM-frequency as a cue to auditory speed perception, thereby revealing an intriguing similarity between auditory and both visual and tactile motion perception. Specifically, we hypothesize that temporal frequency might induce the same patterns of motion illusions and aftereffects in both vision and audition: such a finding would strongly suggest that the different modalities share similar neural mechanisms for motion processing.

2. Results and discussion

We generated rattling sounds by applying sinusoidal amplitude modulations (AM) to pink-noise carrier signals. Sounds were played via a set of eight speakers mounted on a ring with the participants’ head in the centre, thus producing circular motion (figure 1a,b). In a speed-discrimination task, participants reported which of two consecutive moving sounds seemed faster, while across trials we independently varied both the physical speed of the sound source and its AM-frequency. One of the two stimuli presented on each trial—the standard—always moved at the same speed, and had a constant AM-frequency. The other stimulus—the comparison—could move at one of 10 different speeds and have one of 10 different AM-frequencies. The standard and the comparison stimulus were presented sequentially in a randomized order. We fitted a generalized linear mixed model to the proportion of ‘comparison faster’ responses, with speed and AM-frequency as fixed effects. Results showed that perceived speed was influenced by both the sound’s actual speed (z = 4.85; p < 0.0001) and AM-frequency (z = 2.57; p = 0.01; figure 1c, upper panel). That is, sounds appeared to move faster not only with increasing speed, but also with increasing AM-frequency. The relative weight assigned to the two cues (speed and AM-frequency) can be estimated from the orientation of the ‘line of subjective equality’, that is, the combination of speed and AM-frequency at which the standard and comparison stimuli appeared equally fast (figure 1c, continuous line). Results indicate that—on average—participants relied jointly on both cues (figure 1c, upper and lower panels). A control experiment demonstrated that AM-frequency affected perceived speed even when participants were explicitly instructed to ignore such information, and to focus only on the actual speed of the sound (see electronic supplementary material, figure S1, and Methods).

Still, to draw stronger conclusions, we need to ensure that these findings reflect genuine perceptual effects rather than response biases. To this end, we conducted a second experiment in which we investigated whether AM-frequency modulates the auditory motion aftereffect (aMAE). The aMAE is a phenomenon whereby, after exposure to a sound moving in one direction, a subsequent stationary (or slowly-moving) sound appears to move in the opposite direction [16,17]. Motion aftereffects are typically interpreted as adaptation of specialized motion analysers, and their magnitude is proportional to the perceived velocity of the adapting stimulus [16]. Therefore, if there are dedicated neural mechanisms for auditory motion perception that are sensitive to the AM-frequency of sounds, not only speed but also AM-frequency should modulate the aMAE.

In Experiment 2 we induced motion adaptation using a block design. In each block the adapting sound rotated either clockwise or counterclockwise at a constant speed, and it was presented at one of three AM-frequencies. After adaptation, participants reported the direction of motion of a slowly-moving test sound (unmodulated pink noise), which varied in speed across trials to estimate the velocity that was necessary to null the perceived speed of the motion aftereffect (Point of Subjective Stationarity—PSS). To keep participants in an adapted state, top-up adaptation sounds alternated with test stimuli (figure 1d, upper panel). To estimate the PSS (see [16] for a similar procedure to measure the magnitude of the aMAE) we fitted cumulative Gaussian distributions to the proportion of ‘leftward’ (i.e. counterclockwise) responses separately for each participant and adapting stimulus. Following adaptation, the PSS showed a bias in the same direction as the adapting stimulus, meaning that a stationary sound was heard as moving in the direction opposite to the adapting stimulus (figure 1d, lower panel). Importantly, the magnitude of the aMAE increased linearly with the AM-frequency of the adapting stimulus (r² = 0.98), demonstrating that AM-frequency truly affected perceived speed of auditory motion. Moreover, to assess the overall effect of AM-frequency on the aMAE, irrespective of the specific direction of motion, we calculated the clockwise minus counterclockwise difference in PSS for each AM-frequency (figure 1d, inlaid plot). Such difference in the PSS within each AM-frequency showed that, irrespective of the direction of motion, the aMAE was linearly modulated by the AM-frequency: the higher the AM-frequency the higher the aMAE (r² = 0.944).

The dependency of auditory speed perception on the sound’s AM-frequency offers clear ecological advantages because the AM-frequency of sounds produced by moving objects seems to scale with speed (in otherwise isotropic environments). This renders acoustic AM-frequency a relative cue for speed, so that differences in AM-frequency of a moving sound can be directly interpreted as differences in speed. Many aspects of the perceptual system rely on relative cues, such as for example when combining different cues for visual depth perception [18]. In order to make efficient use of such relative cues, the perceptual system exploits natural scene statistics [1,3,6,8,19] to scale such cues and to infer the actual state of the world. The present study demonstrates that the brain successfully capitalizes on an environmental property—the AM-frequency of moving rattling sounds—to perceive auditory speed. Thus, the human auditory system seems to be tuned to the statistical regularities of natural auditory scenes (see also [3]).

The present results tightly mirror previous findings in motion vision, where the spatio-temporal frequency of moving images systematically affects perceived speed: displays having higher temporal frequencies are perceived as moving faster [11], and induce stronger MAE [12]. While in vision these effects seem to depend on the properties of motion detectors, the existence of analogous detectors in audition is still debated (e.g. [20]). The current results show that, just like visual motion detectors [12], the neural circuitry
underlying auditory motion perception is sensitive to the temporal frequency (here the AM-frequency) of the stimulus. Mounting neurophysiological evidence seems to demonstrate that, following the principle of neural economy, the nervous system may detect motion by implementing a small set of canonical computations, which operate in an analogous fashion across the different senses. For instance, temporal frequency also affects tactile speed perception [13–15]. The analogy between our results and previous findings in vision and touch clearly suggests that similar mechanism for motion perception might be at play in different modalities.

3. Methods

(a) Experiment 1: amplitude modulation frequency modulates auditory speed perception

(i) Participants
Nine participants (three males, age: mean ± s.d. = 26.8 ± 4 years), with normal or corrected-to-normal vision and normal audition took part in the study. They provided written informed consent and received 6 euros per hour in return for their participation. The study was carried out in accordance with the Declaration of Helsinki and was approved by the University of Bielefeld ethics committee.

Figure 1. Experimental set-up, stimuli and results. (a) Apparatus. A set of eight loudspeakers was arranged on a circular array around participants’ head. Sound motion was generated by cross-fading the acoustic signals across neighbouring speakers. (b) Stimuli. The amplitude of pink noise signals was modulated by a sinusoidal function. (c) Results of Experiment 1. The upper panel shows the results averaged across participants. The x-axis represents AM-frequency, while the y-axis represents speed. The colour of each cell represents the probability of responding ‘comparison faster’, with warmer colours representing higher probabilities. The dot at the centre represents the standard stimulus. The comparison stimulus varied in AM-frequency (10 steps) and speed (10 steps) according to the method of constant stimuli. The solid line represents the line of subjective equality (LSE), while the dashed lines represent 25% and 75% probabilities of responding ‘comparison faster’. The probability to respond ‘comparison faster’ is higher for both higher speeds and higher AM-frequencies. The lower panel represents the relative weight assigned to AM-frequency by each participant (black bar; grey bar = average ± s.e.m.; see electronic supplementary material, figure S1 for results of the control experiment). Values higher than zero indicate that the participant relies on the information provided by the AM-frequency to solve the task. (d) Experiment 2. The upper panel shows the procedure time line. After 90 s adaptation to a rotating sound, participants indicated the direction of motion of a test stimulus. Top-up adaptation trials alternated with test trials. The lower panel shows the results. The Points of Subjective Stationarity (PSS) are plotted against the AM-frequency and direction of the adapting stimulus (lower panel). Positive values indicate clockwise motion. The magnitude of the aMAE increases linearly with the AM-frequency of the adapting stimulus. The inlaid plot represents the clockwise minus counterclockwise difference in PSS for each AM-frequency. (Online version in colour.)
(ii) Apparatus

Eight loudspeakers were equally spaced along a circle (diameter = 91.5 cm). Each speaker consisted of two drivers, placed at a distance of 32 cm on the vertical axis, and mounted on an octagonal frame (figure 1r; see also [19] for a similar apparatus and procedure). Participants sat in the middle of the circle, with their ears at approximately the height of the sound source. Participants wore a head-mounted display (Zeiss 3D visor, eMagin, Bellevue, WA, USA), which presented the experimental instructions, while hiding the setup from view. To stabilize participants’ head, we used a helmet that was mounted on a suspended pole attached to the roof. This guaranteed that participants maintained their head still and at the right height in the middle of the circle.

(iii) Stimuli and procedure

To generate moving rattling sounds, stimuli were cross-faded between neighbouring loudspeakers (see e.g. [21]). Auditory carrier signals consisted of 1/3-octave (pink-noise, sampling frequency 44.1 kHz), whose amplitude was modulated by a sinusoidal function (i.e. the modulating signal, see below and figure 1b). Stimulus presentation was controlled by custom software that was based on the Psychotoolbox 3.0 [22] running on a computer equipped with an M-Audio Delta 1010LT PCI audio card. Sounds were amplified by a Stageline-IMG STA-1508 amplifier and played at an average intensity of 66 dB SPL. They were smoothly ramped on and off on a window of 0.5 s. The intensity of the sounds produced by the eight speakers was carefully equalized using an SPL meter. Moreover, the intensity of each speaker was slightly jittered on a trial-by-trial basis by multiplying the signals sent to each speaker by a random factor between ± 0.05. This was done to prevent participants from using minor differences in loudness across speaker as additional cues.

On each trial, two consecutive sounds were presented with an inter-stimulus interval of 0.5 s. Participants had to judge whether the first or the second stimulus seemed to move faster by pressing the left or the right mouse button, respectively. Each pair of stimuli consisted of a standard and a comparison stimulus, played in pseudo-randomized order. Within each trial, one of the two stimuli—namely the standard stimulus—moved at a constant speed of 270°/s, and had an AM-frequency of 10.61 Hz. The other stimulus—the comparison—could move at one of 10 different speeds, logarithmically spaced between 180 and 405°/s (i.e. 180, 197, 215.5, 235.9, 258.12, 282.48, 309.1, 338.22, 370.14, 405°/s), and could have one of ten different AM-frequencies, logarithmically spaced between 7.5 and 15 Hz (i.e. 7.5, 8.1, 8.75, 9.45, 10.21, 11.02, 11.91, 12.86, 13.89, 15 Hz; see electronic supplementary material, audio file ‘Stimuli’). Thus, the combination of all speeds and AM-frequencies determined a 10 × 10 grid, over which we could fit a 2D psychometric function (a psychometric surface).

On each trial, both standard and comparison stimuli moved in the same direction (clockwise or counterclockwise). The direction of rotation varied pseudo-randomly across trials. At the beginning of each trial, and for its entire duration, an arrow was presented at the centre of the head-mounted display, to inform participants on the direction of motion of the pair of stimuli in that trial. In order to prevent participants from using stimulus displacement (i.e. the length of the trajectory covered by the moving sound) as a cue for speed, we randomly varied the duration of each stimulus between 1.5 and 2.5 s (e.g. see [23]). Similarly, we also randomly varied the starting position of the sound source across stimulus presentations (cf. [23]). The range of speeds was chosen based on preliminary observations and previous studies [19,24] demonstrating that rotational motion within this range can be reliably perceived. After the experimental session, all participants were asked to verbally report whether they could perceive the sounds as rotating around them. They all reported a clear perception of rotational motion throughout the entire experiment. Overall, the experiment consisted of two sessions, taking place in 2 consecutive days, and lasting approximately one hour and a half each, for a total of 800 trials. To motivate participants, every 25 trials the head-mounted display showed the percentage of correct responses and participants could take a break. Such a feedback was not displayed on a trial-by-trial basis to prevent learning.

Before the first experimental session, participants underwent a training session (80 trials) to familiarize with the stimuli and task. During the training, an auditory feedback (a beep) was provided whenever participants produced an incorrect response. Both the standard and the comparison had the same AM-frequency of the standard stimulus used in the main experiment (i.e. 10.61 Hz).

(iv) Data analyses

For each participant and each combination of speed and AM-frequency, we calculated the proportion of ‘comparison faster’ responses. The effects of actual speed and AM-frequency of moving sounds on speed perception was assessed by fitting the probability of responding ‘comparison faster’ with a generalized linear mixed model (GLMM [25,26]), with speed and AM-frequency as fixed effects. Given that discrimination performance is at chance level when standard and comparison are physically identical, we constrained the intercept of the psychometric function to reduce the number of fitted parameters. To account for the heterogeneity among different subjects, we included the random intercept, the slope, and their interaction as random effects parameters. A Probit link function was applied and maximum likelihood (ML) estimation was used to estimate the parameters. We tested the significance of each parameter with the Wald test. The analysis was performed using R and the MERpsychophysics toolbox developed by [26].

(v) Results

Results showed that the perceived speed of rattling moving sounds is influenced by both the actual speed of the sounds and their AM-frequency (see Main Text and figure 1c, upper panel). To test whether speed and AM-frequency interacted with each other in modulating participants’ responses, we fitted another GLMM including also the speed by frequency interaction and compared this model with the previous one. According to the Akaike Information Criterion (AIC, [27]), the former model fitted the data better than the second one, indicating a lack of interaction between the two cues. In the present 2-factorial analysis (i.e. involving speed and AM-frequency), the point of subjective equality (PSE, i.e. the combinations of speed and AM-frequency at which the standard and the comparison stimuli appeared equally fast) is represented as a line: the line of subjective equality (LSE, figure 1c, upper panel, solid line). The orientation of the line represents the relative weight assigned to the two cues for auditory speed perception. A horizontal orientation would indicate that participants rely only on the sounds’ actual speed to estimate speed perception, while a vertical orientation would show that participants take into consideration only AM-frequency when judging the speed. The diagonal orientation of the line found in Experiment 1 indicates that participants rely on both cues for auditory speed perception (figure 1c, upper panel).

To investigate the influence of AM-frequency on speed perception at the level of the individual observers, we estimated the weight that each participant has given to AM-frequency
\[ a_{\text{AM freq}} = \frac{\beta_{\text{AM freq}}}{\beta_{\text{AM freq}} + \beta_{\text{speed}}} \]

Here \( \beta_{\text{AM freq}} \) represents the linear coefficient of AM-frequency and \( \beta_{\text{speed}} \) the coefficient of speed. Values higher than zero indicate that the participant relies on the information provided by AM-frequency to solve the task. Individual data (figure 1c, lower panel) show that, although the weight given to AM-frequency differs across participants, most of them significantly relied on AM-frequency when reporting the sounds’ perceived speed (\( t \)-test against zero, \( t_s = 3.4, p = 0.009 \)) (see electronic supplementary material, data analyses for more details).

(b) Experiment 1—control: the influence of task instructions

In order to rule out the possibility for response strategies—instead of perceptual effects—dominating results, we ran a control experiment. Specifically, we investigated whether participants might be able to ignore the information provided by AM-frequency, when explicitly instructed to do so, or whether such a cue is automatically integrated with the other speed signals. Participants were presented with the same stimuli and task of the main experiment. The experimenter explained how the stimuli were created and asked participants to give the response by focusing on the actual speed of the sounds only (i.e. the speed of the spatial displacement of the sound), while ignoring the information related to AM-frequency.

(i) Participants

Five naive participants (1 male, \( M = 26.6, \) s.d. = 4.4), with normal or corrected-to-normal vision and normal audition took part in the study. None of them had taken part in the previous experiment.

(ii) Apparatus, stimuli, procedure and data analyses

Procedures, behavioural measures and analyses were the same as in the previous experiment. Prior to the beginning of the experimental session, the experimenter explained participants in detail how stimuli were created and specified that the sounds could independently vary over both, speed of motion and AM-frequency. To give a practical demonstration, the experimenter played pink noise alone, then played different sounds whose amplitude was modulated by different AM-frequencies. Participants were told that speed and AM-frequency were independently varied and that, the AM-frequency of the sound was task-irrelevant and uninformative as to with which speed the stimulus was actually rotating. Thus, they were explicitly instructed to focus only on the actual speed of the stimuli while performing the task and to ignore the AM-frequency.

(iii) Results

Like in the main experiment (see main text), the results showed that both the actual speed and the AM-frequency of the sounds affect the perceived speed of moving sounds (see electronic supplementary material, figure S1, upper panel), with a higher probability of responding ‘comparison faster’ for combinations of higher speeds and AM-frequencies. As in the previous experiment, the model that best fitted the data was the one including the fixed effects of speed and AM-frequency only, without the speed by frequency interaction, and including the random intercept, slope, and interaction as random effects.

Compared to the previous experiment, the LSE (electronic supplementary material, figure S1, upper panel) in the current control experiment is more tilted toward the horizontal axis.

This demonstrates that, although overall speed is weighted more in this experiment than in the previous one, speed perception was nevertheless affected by AM-frequency in all participants, as confirmed by the fact that the weights given by each participant to AM-frequency were higher than zero (\( t_4 = 2.96, p = 0.04 \)) (electronic supplementary material, figure S1, lower panel). Put in other words, AM-frequency could not be fully ignored and was automatically integrated to form a combined percept of auditory speed. In both experiments we found large individual differences in weighting of speed and AM-frequency, with some participants being almost completely dominated by the AM-frequency cue, while others almost ignored it. It might be the case that the reliability of such sensory cues differs across participants. That is, while some observers might be particularly sensitive to speed, others might find it hard to detect speed, and thus rely only on the AM-frequency cue to solve the task. This would be in accordance with recent findings highlighting large individual differences in sensitivity, as well as in prior expectations (e.g. \( [28,29] \)).

(c) Experiment 2: amplitude modulation frequency influences auditory motion aftereffect

(i) Participants

Seven naive participants (3 males, \( M = 22, \) s.d. = 0.9), with normal or corrected-to-normal vision and normal audition took part in the study. None of them took part in the previous experiments.

(ii) Apparatus

The apparatus was the same as in the previous experiment.

(iii) Stimuli and procedure

For an even stronger test of AM-frequency affecting perceived auditory motion we used the indirect measure of the motion aftereffect. We used this indirect measure because it is unlikely to be affected by any form of response bias and thus should reveal solely perceptual effects. To this end, we played rattling sounds rotating around the participant’s head at a constant speed and investigated whether the AM-frequency of a sound affects the auditory motion aftereffect (aMAE) following exposure to that sound.

The experiment consisted of six blocks. Each block started with an adaptation phase where a sound continuously rotated around participants’ head for 90 s (figure 1d, upper panel). Participants were then presented with a test stimulus (1 s), slowly moving either clockwise or counterclockwise. They were then asked to report its direction of motion (i.e. rightward or leftward, respectively) by pressing one of two mouse buttons. In order to keep participants in an adapted state, they received top-up adaptation stimulation for 5 s between two consecutive trials. The test stimulus started 0.5 s after the top-up adaptation (figure 1d, upper panel). The adapting stimulus moved always at the same speed as the standard stimulus used in the previous experiments (i.e. 270°/s). Within each block, the adapting stimulus could have one of three AM-frequencies, logarithmically spaced between 9 and 15 Hz, and one motion direction (clockwise or counterclockwise), thus yielding to six combinations of AM-frequencies and motion directions (i.e. ±9°, ±11.62°, ±15 Hz, where positive values indicate clockwise motion). The test stimulus consisted of pink noise with no amplitude modulation.

On each trial, the sound could randomly move either clockwise or counterclockwise, and start from one of five possible start positions (i.e. 0°, ±22.5°, ±45°), where 0° indicates participant’s body midline. On each trial the speed of the test stimulus was determined using a custom adaptive procedure based on the QUEST [30]. Specifically, we used two QUESTs
running in parallel (one with a starting velocity of \(-5\, \text{cm/s}\), the
other with a starting velocity of \(5\, \text{cm/s}\)). In a classic QUEST pro-
cedure, the psychometric function is estimated iteratively after
each trial, and the stimulus used on each trial is determined by
the estimated point of subjective equality (in our case, the
Point of Subjective Stationarity—PSS). This method, however,
does not provide enough observations away from the PSS,
making it hard to estimate the slope of the psychometric func-
tion. Moreover, given that most stimuli are close to the PSS (i.e.
where response is at chance), participants often find the task
too hard and frustrating. Therefore, the speed of the stimuli
used on each trial was determined by the PSS estimated by the
QUEST, to which we added a random value taken from a Gauss-
ian distribution (with a standard deviation equal to one JND, as
determined in pilot observations). By presenting also easier
trials, such a procedure is more enjoyable by participants, and
it also allows to reliably estimate the slope of the psychometric
functions. This procedure was selected after preliminary obser-
vations indicating a high variability in participants’ ability to
correctly identifying motion direction, that would have made it
hard to select an effective set of velocities, as required by the
method of constant stimuli. Overall, the experiment consisted
of two sessions of three blocks each, taking place in 2 consecutive
days, and lasting approximately one hour each. Each block (i.e.
each combination of AM-frequency and motion direction) con-
sisted of 80 trials, for a total of 480 trials. In order to minimize
any order effect, the order of the six blocks was randomized
across participants. A few minutes break was set between two
consecutive blocks. Before the first block, participants were
required to perform a quick training (80 trials, without adap-
tation) in order to familiarize with the stimuli and task.
Participants were presented with a single pink noise sounds rando-
moly moving either clockwise or counterclockwise at one
of four different velocities, linearly spaced between \(+5\, \text{cm/s}\)
(i.e. \(\pm 1, \pm 2.34, \pm 3.67, \pm 5\, \text{cm/s}\)). Their task was to indicate the
direction of motion. During the training, an auditory feedback
(a beep) was provided after incorrect responses.

(iv) Data analyses
We fitted cumulative Gaussian distributions to the proportion of
‘leftward’ (i.e. counterclockwise) responses [31] for each of the six
combinations of motion direction and AM-frequency of the
adapting stimulus. This procedure was separately performed
for each participant. As a measure of the magnitude of the
aMAE, we calculated the velocity at which the test stimulus
appeared stationary to the participants (i.e. the Point of Subjec-
tive Stationarity, PSS; see [16,32] for a similar approach). If
there is an aMAE, the PSS should shift toward (negative afteref-
fect) or away from (positive aftereffect) the adapting stimulus.

Data accessibility. Datasets supporting this article can be found in Dryad repository:
http://dx.doi.org/10.5061/dryad.62p91 [33].

Authors’ contributions. Designed the experiment: I.S., C.V.P., M.O.E.; ana-
lysed the data: I.S., C.V.P.; performed the experiment: I.S.; wrote the
paper: I.S., C.V.P.; M.O.E.

Competing interests. The authors have declared that no competing inter-
ests exist.

Funding. I.S. was supported by the Cluster of Excellence Cognitive Interaction Technology ‘CITEC’ (EXC 277) at Bielefeld University,
which is funded by the German Research Foundation (DFG), project ICSPACE. C.V.P. was supported by the 7th Framework Programme
European Project ‘Wearpac’ (601165).

Acknowledgements. We thank Loe C.J. van Dam and Alessandro Mosca-
telli for helpful comments on an early version of this manuscript. We
further thank Mario Botsch and Eugen Dyck for contributing to the
experimental set-up, and Anna Oppenborn for assistance in collect-
ing the data of the last experiment.

References

8. Ernst MO. 2007 Learning to integrate arbitrary signals from vision and touch. J. Vis. 7, 1–14. (doi:10.1167/7.5.7)


Supplementary information

Figure S1. Results of the control experiment (see Figure 1). Participants were presented with the same stimuli and task of Experiment 1 and were explicitly instructed to actively ignore the information provided by the AM-frequency. The upper panel shows the results from all participants aggregated. The x-axis represents AM-frequency, while the y-axis represents speed. The color of each cell represents the probability of responding “comparison faster”, with warmer colors representing higher probabilities. The dot at the center shows the standard stimulus. The solid line represents the line of subjective equality (LSE, i.e. the combination of speed and AM-frequency at which the standard and comparison stimuli appear equally fast), while the dashed lines represent 25% and 75% probabilities of responding “comparison faster”. The lower panel shows the scaled weight of the AM-frequency for each participant (black bars), and the average (grey bar). Values higher than zero indicate that the participant relies on the information provided by the AM-frequency to solve the task. The bar represents the standard error of the mean.

Supplementary data analyses

For each combination of speed and AM-frequency, we calculated the proportion of “comparison faster” responses. Given that we assessed perceived speed as a function of two independent variables (speed and AM-frequency), psychophysical data can be represented as psychometric surfaces (instead of psychometric curves, which are better suited to represent perception as a function of a single independent variable, see [S1] and [S2] for a similar approach). Therefore, the probability of responding “comparison faster” as a function of speed (Y-axis) and AM-frequency (X-axis) is plotted in Figure 1 of the main text and in Figure S1 with warmer colors representing higher probabilities. Such a surface can be read vertically (i.e., column-wise) and horizontally (i.e., row-wise). The X-axis
represents the 10 AM-frequencies used in the experiments, while the Y-axis represents each of the 10 speed values. Thus, reading each column vertically allows appreciating the effect of speed for each single AM-frequency value. Likewise, reading each row horizontally allows seeing the effect of the different AM-frequencies for each speed.

Data were analyzed by fitting the probability of responding “comparison faster” with a generalized linear mixed model (GLMM [S3-S4]). Traditional approaches in psychophysics generally apply two-level analyses: first, the parameters of the psychometric functions (i.e., the point—or in our case the line-of subjective equality and the just noticeable difference) are estimated for each participant and condition, and then a second (inferential) analysis (e.g., ANOVAs, t-tests, etc.) is performed on such estimated parameters [S4]. Compared to two level analyses, the main advantage of the GLMM is that the variance is divided into a fixed and a random component. The fixed component tests the effect of the independent variables manipulated in the study (in our case the AM-frequency and the speed), while the random effect accounts for the heterogeneity among different participants. Thus, the GLMM allows for a joint analysis of clustered categorical data at both the population and individual levels, while dramatically reducing the number of fitted parameters. Moreover, such a model allows for an easy assessment of the goodness of fit, and thus the selection of the best statistical model (for instance, via the Akaike Information Criterion [S5], see main text).

To further explore the inter-subject variability (i.e., to investigate the strength of the influence of AM-frequency on speed perception in each participant), we run a second analysis at the individual level, beside group-level analyses. For each participant we fitted the same generalized linear model used for the group analysis, but applied at the individual level (i.e., for each participant we fitted the probability of responding “comparison faster”, with speed and AM-frequency as fixed effects, and by applying a probit link function). Clearly, this implies dropping the random effects, given that such analysis is performed at the single participant level. For each participant, the linear coefficients of AM-frequency and speed ($\beta_{\text{AM freq}}$ and $\beta_{\text{Speed}}$, respectively) were used to calculate the weight given by each participant to the AM-frequency ($\omega_{\text{AM freq}}$), by dividing the linear coefficient of the AM-frequency by the sum of the two linear coefficients (i.e., $\omega_{\text{AM freq}} = \beta_{\text{AM freq}} / (\beta_{\text{AM freq}} + \beta_{\text{Speed}})$). In other words, the AM-frequency coefficient is normalized based on the sum of the AM-frequency and speed coefficients, with values higher than 0 indicating that the participant relies—at least in part—on the AM-frequency cue to estimate the speed of moving sounds.

Supplementary references


