Hearing Research 459 (2025) 109227

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Unveiling the mind's ear: Understanding the science behind auditory processing using illusions

Anusha Yasoda-Mohan [a,b,1], Feifan Chen [a,b,1], Sven Vanneste [a,b,c,*]

[a] Lab for Clinical and Integrative Neuroscience, Trinity College Institute for Neuroscience, Trinity College Dublin, D02 PN40, Dublin, Ireland
[b] School of Psychology, Trinity College Dublin, D02 PN40, Dublin, Ireland
[c] Global Brain Health Institute (GBHI), Trinity College Dublin, D02 PN40, Dublin, Ireland

ARTICLE INFO

Keywords:
Auditory processing
Spectral
Temporal
Spatial
Multisensory
Applications of illusions

ABSTRACT

Our perceptual experience is not purely driven by the information our senses receive but is an active combination of prior experience and the sensory information that we receive. This is promisingly demonstrated by an illusory experience. Particularly in the auditory domain, we encounter illusions that serve as a tool to understand the more holistic and bigger umbrella of auditory processing. In the current manuscript, we review spectral, temporal, spatial auditory and multisensory processing through the lens of auditory illusions. We review examples of auditory illusions that could serve as models to better comprehend these sub-domains of auditory processing. We also explore the literature where these illusions have been used as a causal human model to unravel mechanisms behind neuropathology and conclude with futuristic applications of auditory illusions.

1. Introduction

We hear the world so effortlessly, but the intricate mechanisms that underlie this process is still poorly understood. The apparently coherent and seamless perceptual experience must be assembled from the discontinuous and ambiguous information reaching our senses. This way, auditory perception is a complex process through which the brain interprets sounds by analysing various features including spectral, temporal and binaural information to recognise auditory objects, predict when an auditory event is happening, and identify the source of the auditory object respectively (Oxenham, 2018). Notably, each of these dimensions plays a crucial role in how we make sense of the sounds around us.

The most fundamental auditory process is to discriminate objects based on its spectral profile. As the complexity of this profile increases, the ability to process fundamental frequencies becomes key in comprehending speech, music and environmental context (Oxenham, 2012). In music, sequences of pitches create melodies, while harmonies arise from simultaneous pitch combinations. In spoken language, pitch variations establish prosody, and in tonal languages like Mandarin and Cantonese, pitch contours convey distinct word meanings. In complex auditory environments, pitch differences help listeners distinguish and interpret overlapping sound sources (Oxenham, 2012). Temporal processing forms the foundation for capturing the timing of sounds, which is essential for tracking rhythm in speech and music. This ability allows us to differentiate between rapid and slower sequences of sounds (Bizley and Cohen, 2013) which is important for the categorisation in auditory perception. Parallelly, temporal integration and continuity is also key to assimilating categorical components into a seamless perception of continuous speech and melody (Jain and Nataraja, 2019). This processing of auditory categorisation and integration in its temporal structure is crucial in parsing information into the auditory sensory memory (Rimmele et al., 2015). Auditory spatial processing enables the determination of a sound source's location, which is vital for auditory stream segregation. The brain accomplishes this by analysing intra-aural time differences—the difference in time it takes for a sound to reach each ear—and interaural level differences, which involve variations in perceived loudness between the ears (Shamma, 2001).

In addition to processing spectral, temporal and spatial information, the ability to integrate sensory inputs—known as multisensory integration—is crucial for effective auditory perception and cognitive functioning (Stein and Stanford, 2008). Multisensory information is combined at various processing levels to enhance stimulus detection, localization, and response accuracy, as well as to accelerate reaction

* Corresponding author at: Lab for Clinical and Integrative Neuroscience, Trinity College Institute for Neuroscience, Trinity College Dublin, D02 PN40, Dublin, Ireland.
E-mail address: sven.vanneste@tcd.ie (S. Vanneste).
[1] Both authors contributed equally to the project.

https://doi.org/10.1016/j.heares.2025.109227
Received 31 October 2024; Received in revised form 6 February 2025; Accepted 24 February 2025
Available online 25 February 2025

times (Wallace et al., 2020). At the lower processing levels, the physical characteristics of stimuli from different sensory modalities are aligned, ensuring spatial, temporal, and intensity coherence (Stein and Meredith, 1993). In contrast, higher-order multisensory integration involves the interaction and mutual influence of different sensory modalities, which is essential for adaptive behaviour (Wallace et al., 2020).

Recent research involving clinical populations has provided valuable insights into the functioning of these processes and the underlying mechanisms that support them. For instance, disrupted temporal processing is commonly observed in individuals with dyslexia or schizophrenia, affecting the ability to perceive and interpret speech correctly (Carroll et al., 2008; Casini et al., 2018). Patients with stroke or lesions affecting the auditory system or its associated pathways frequently demonstrate impairments in sound localization, offering deeper insights into the mechanisms of spatial segregation and speech perception in noisy environments (Griffiths et al., 1997). On the other hand, hearing loss and cochlear implantation provide a window to study the mechanism of pitch perception and how pitch processing is disrupted, particularly in complex auditory environments like music (Oxenham, 2008).

Alternatively, auditory illusions—experimentally induced perceptual distortions of auditory stimuli—serve as a compelling model for causally investigating the behavioural relevance and the psychophysical characteristics of different auditory processes in the spectral, temporal, spatial and multisensory domains. By exploring these phenomena, we can gain deeper insights into the complexities of auditory processing, from basic perceptual functions to higher-order cognitive mechanisms. In some cases, manipulations of auditory stimuli are sufficient to generate different auditory illusions e.g., auditory continuity illusion (Miller and Licklider, 1950), the Precedence Effect (Wallach et al., 1949), and Zwicker Tone illusion (Zwicker, 1964). Alternatively, auditory illusions can also be perceived as a result of conflicting information from different sensory modalities, such as McGurk effect (McGurk and MacDonald, 1976) and ventriloquist's illusion (Thurlow and Jack, 1973).

Despite extensive research, current studies often treat auditory illusions in isolation, leading to fragmented insights that fail to fully capture their broader implications for auditory processing. This presents a significant gap in the literature when it comes to synthesizing this knowledge and leveraging these phenomena to enhance our understanding of auditory perception. A comprehensive discussion that integrates these individual findings is crucial for constructing a more cohesive understanding of how auditory illusions can inform the mechanisms underlying perception. By unifying these knowledge fragments, we can better appreciate the role of auditory illusions as a powerful tool for modelling and exploring the complexities of auditory perception. This approach not only advances theoretical frameworks but also has practical implications for improving auditory interventions and technologies.

Thus, this review aims to highlight key examples of auditory illusions which have been extensively studied in the fields of psychoacoustics, and to underscore their significant value as models for understanding auditory processing across temporal, spatial, spectral and multisensory domains. Additionally, a narrative review on auditory illusions is timely and relevant, as it consolidates existing knowledge, identifies gaps in current research, and provides a comprehensive framework for future studies, ultimately advancing our understanding of how the auditory system operates under both typical and atypical conditions.

2. Understanding auditory spectral domain processing through illusions

The most basic function of the auditory system is to recognise sounds based on its frequency profiles. Each sound – a man/woman/child's voice, the growl of an animal or the tune of a song has a very specific spectral profile and has both an ecological and evolutionary validity which triggers a particular kind of behavioural response. The perceptual correlate of the frequency of the sound or "pitch" depends on where the sound vibrations are encoded on the basilar membrane in the inner ear and the amplitude of these vibrations encodes the perceptual correlate of the intensity of the sound or "loudness" (Frank and Jane, 2020).

2.1. Understanding the tonotopic pitch organisation of the auditory system

A purely spectral illusion that utilises this frequency-specific organisation of the inner ear is the *Zwicker Tone illusion* (Fig. 1a). The Zwicker Tone illusion is an auditory after-image induced following the presentation of a notched noise (i.e. broadband noise with a spectral gap) (Zwicker, 1964). The width of the notch and the frequency of the missing centre entails an essential characteristic and success of the afterimage perception (Zwicker, 1964). From previous studies, it is shown that a central frequency between 4000 – 6000 Hz and a notch of width 0.75 – 1 octave may be ideal to reliably produce a Zwicker Tone illusion (Leske et al., 2014; Mohan et al., 2020; Norena et al., 2000). The Zwicker Tone illusion highlights the role of lateral inhibition in auditory processing (Norena et al., 2000; Pantev et al., 1999; Rhode and Greenberg, 1994). When a band of frequencies is suppressed, the excitation of the neurons tuned to neighbouring frequencies tend to "fill-in" for the missing information (Franosch et al., 2003). Recent studies showed a varied susceptibility of Zwicker Tone illusion across subjects even when using the optimal noise characteristics (Leske et al., 2014; Mohan et al., 2020), suggesting an involvement of other mechanisms during the perception of the illusion (Hullfish et al., 2019; Schilling et al., 2021). Furthermore, there is an ongoing debate as to whether the Zwicker Tone illusion is induced by bottom-up peripheral processes or also accompany top-down central processes in its generation (Leske et al., 2014; Mohan et al., 2020; Noreña and Eggermont, 2003). Given these different characteristics, the Zwicker Tone illusion is proposed to be a human model for tinnitus (Norena et al., 2000) – the perception of a continuous tone in the ear in the absence of an external stimulus (Baguley et al., 2013). Thus, the Zwicker Tone illusion illustrates the brain's ability to generate perceptual experiences that extend beyond the immediate physical stimulus, highlighting the significant role of higher-level processing and compensatory mechanisms in shaping auditory perception.

2.2. Comprehending the complexity of understanding natural sounds

However, natural sounds including speech, music and vocalisations do not contain one frequency but are organised in harmonics where frequency components are multiples of a common fundamental frequency (f_0). By controlling the phase and amplitude of these different frequency components, one can modulate the interacting effects of the fundamental frequency with its harmonics to modulate the perception of the pitch thereby producing three illusions – *periodicity pitch, circular pitch and tritone paradox illusions.*

2.2.1. Periodicity pitch or missing fundamental illusion

Periodicity pitch or *missing fundamental illusion* refers to the auditory illusory phenomenon where the perceived pitch of a sound remains the same even when the f_0 fundamental frequency is absent or masked (Licklider, 1954) (Fig. 1b). To induce the missing fundamental illusion, complex harmonic tones are first created with the f_0 removed from the complex tone, leaving only the higher harmonics (e.g., f_1, f_2). Interestingly, when exposed to the manipulated tone, participants are still able to perceive the fundamental pitch as if it was present, even in its physical absence. When the tone is played, although there is a lack of energy at f_0, the combination of the other harmonics creates a series of periodic peaks whose frequency corresponds to f_0 (Chialvo, 2003; Plack et al., 2005). This phenomenon is hypothesised to occur based on the brain's ability to the infer the fundamental frequency from the resolved harmonics of the sound, which are multiples of the fundamental frequency (Oxenham, 2012; Plack et al., 2005). Furthermore, the tonotopic organization of the auditory cortex is based on pitch rather than frequency and therefore

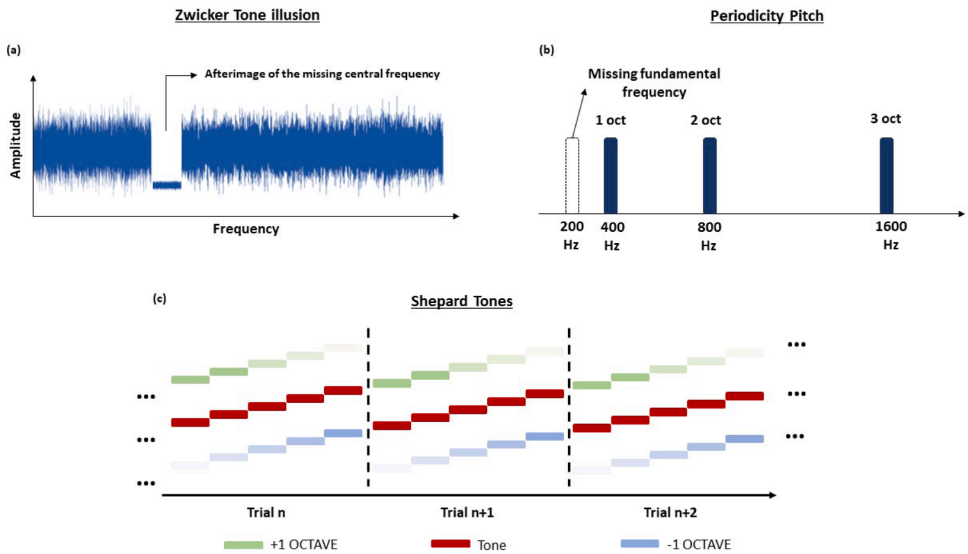


Fig. 1. Examples of auditory spectral-based illusions. (a) The Zwicker Tone illusion, where a broadband spectrum with a gap around a central frequency produces an afterimage of the missing frequency, (b) Periodicity pitch, where the deleted fundamental frequency is still perceived when the higher harmonics are present in a complex sound, (c) Shepard Tones, the illusion of an ever increasing or decreasing pitch generated by 12 tones that are a semitone apart.

this signal periodicity is transmitted to the brain (Pantev et al., 1989). A direct application of the periodicity pitch illusion is the use of analogue telephone lines which typically transmit frequencies between 300 – 3000 Hz, thereby cutting off fundamental frequencies of human voice between 125 – 200 Hz. Nevertheless, we are able to identify the talker and hold seamless conversations.

2.2.2. *Circular pitch illusion (Shepard tones) and the tritone paradox*

The *circular pitch illusion* is particularly used in Western music. Based on the 12-tone sequences, Roger Shepard and James Tenney developed what are called *Shepard tones* or Shepard staircase (Fig. 1c). Shepard tones are a special type of complex sound consisting of 12 tones spanning 1 octave, each of which is a semitone apart (i.e. differ by 6 % in fundamental frequency) (Shepard, 1964). This, described by Shepard in 1964, creates an auditory illusion of a tone that seemingly ascends or descends endlessly with a distinctive timbre, without it physically becoming higher or lower (Shepard, 1964). This is a “musical” analogous to the spinning barber pole where the stripes appear to be traveling up or down the length of the pole. The amplitudes of individual Shepard tones is determined by a Gaussian spectral envelope distributed on a logarithmic frequency scale, through which middle frequencies are emphasised while gradually decreasing the volume of extremely high and low frequencies. However, since the spectral envelope maintains a consistent shape, when the component frequencies shift beyond the audible range, they are replaced by new components at the lower end of the spectrum, effectively creating a seamless loop. This illusion can be reversed to perceive a decreasing pitch when the tones are circled the other way (Shepard, 1964). This complex auditory perceptual model has inspired numerous studies to investigate the active and predictive nature of auditory perception (Chambers et al., 2017; Chambers and Pressnitzer, 2014). Rather than passively registering incoming acoustic information, the brain utilizes prior context, expectations of continuity, and adaptive mechanisms to shape its interpretation of sound, even for basic perceptual features like pitch. The *tritone paradox* is a variation of the Shepard tones, wherein the change in fundamental frequencies between each Shepard tone is half an octave, rather than 1/12 of an octave (Scharine and Letowski, 2009).

3. Understanding auditory temporal domain processing using illusions

Recognition of sounds from an evolutionary standpoint or comprehending speech and music is not just a spectral phenomenon. Timing of when the different sounds occur is key to segregating and passing them into memory or assimilating them into one seamless perceptual experience. Temporal processing is one of the most fundamental functionalities of the auditory system. Temporal processing involves a multitude of sub-domains (Shinn, 2003) including (i) temporal ordering or sequencing – the processing of multiple auditory stimuli in the order of their occurrence, (ii) temporal discrimination or resolution – this refers to the shortest time in which a person can discriminate between two stimuli, (iii) temporal integration – the time period in which the auditory system assimilates the sound energy and affects the loudness perception of the sound based on its duration and (iv) temporal masking – the ability of an auditory stimulus to change the perception of other sounds that precede or succeed it.

3.1. *Exploring temporal discrimination and sequencing using illusions*

The temporal discrimination threshold (TDT) is the minimum time required by a person to perceive two sounds or two stimuli from different sensory modalities as asynchronous – i.e. not occurring at the same time (Artieda et al., 1992; Conte et al., 2017). This is also called “simultaneity judgement” and works together with the ability to temporally order different sounds or stimuli from different sensory domains. TDT is shown to be shorter in women <40 years, however, it is shown to decline three times faster in women than in men after 45 years (Mc Govern et al., 2017; Williams et al., 2015). TDT also changes in different disorders of the basal ganglia (Conte et al., 2017). Interestingly, TDT is smaller in musicians compared to non-musicians and also declines slower in musicians with age (Rammsayer et al., 2012). This means that musicians are able to discriminate sounds that occur much closer in time compared to non-musicians.

3.1.1. *Time shrinking and time stretching illusions*

Temporal discrimination and ordering become crucial in the categorical perception of speech and music. Temporal discrimination can be modulated when sounds are presented in a stream of stimuli and separated by temporal gaps. Short time gaps between stimuli ($< \sim 200$ ms) are considerably underestimated when preceded by shorter time gaps – i.e. when three stimuli are presented in a sequence, the perception of the time gap between sound 2 and sound 3 (t_2) is significantly underestimated when the time gap between sound 1 and sound 2 (t_1) is lesser than t_2. This phenomenon is called *time shrinking* (Nakajima et al., 2004). The crucial factor controlling this illusion was the time difference between t_1 and t_2 when it is ~100 ms. When a model was fit to understand the dynamics of the time shrinking illusion, it was found that t_2 was assimilating on top of t_1. In contrast, an overestimation of the second interval was achieved when $t_2 - t_1 > 240$ms. This was termed *time stretching illusion* (Nakajima et al., 2004). In fact, the two-time intervals could even be perceived as equal or isochronous when it satisfied the criteria -80 ms $\leq t_1 - t_2 \leq 40$ms. Time stretching could also be observed when a sinusoidal wave was followed by a noise stimulus, resulting in an overestimation of the duration of the sine wave (Sasaki et al., 2010) (Fig. 2a).

3.1.2. Isochronicity illusions

Isochronous sounds – or sounds that are equally spaced in time – form the basis of rhythm, both in speech and music (Ravignani and Madison, 2017). Our ability to order sounds in regular, recurring units is the most fundamental principle of Gestalt perception which posits that global perceptual organisation is achieved on the basis of similarity and the spatial and temporal proximity of units (Wertheimer, 1938). The temporal structure of speech and music spontaneously triggers perceptual grouping aiding in efficient processing. It is interesting to note that this rhythmic Gestalt is differently organised in musicians vs non-musicians. A version of the time-stretching illusion was observed in non-musicians when asked to estimate the time gap between two different rhythmic sequences, where musicians performed better, and non-musicians overestimated the difference (Geiser and Gabrieli, 2013). Such an ability to temporally categorise sound sequences is particularly necessary to memorise, imitate or respond quickly to what people are listening in a conversation.

3.2. Understanding temporal masking and integration using illusions

3.2.1. Auditory continuity illusion

As much as categorising auditory stimuli is important, integrating them is equally important. Similar to the visual domain – when pictures are presented faster than the threshold of simultaneity, they are perceived as a continuous motion picture – the auditory domain also produces an illusion of continuous perception of speech and music thereby "blurring" the categorical transition between different syllables. Furthermore, perceptual "filling-in" is a phenomenon that our sensory system uses to actively compensate information that is masked by other competing signals, in order to maintain a coherent perceptual representation of the current state (Komatsu, 2006). In the auditory modality, one illustration of this mechanism is the *auditory continuity illusion* which leverages the ability of the temporal masking in the auditory system (Fig. 2b). During the auditory continuity illusion, an auditory stimulus that would typically be discontinuous due to a temporal gap, can be perceived as continuous when the gap is obscured by a loud masking noise (Bremer et al., 1977; Miller and Licklider, 1950; Warren et al., 1988). This phenomenon, often referred to as "perceptual restoration" (Warren, 1970) or "auditory induction" (Warren, 1970), highlights the auditory system's ability to reconstruct a seamless perceptual experience by integrating the segments preceding and following the occluded signal. Experiments on continuity illusions using non-speech sounds provide a more detailed exploration of the underlying mechanisms. For example, similar illusory phenomena have been experimentally induced with simple tonal stimuli, offering insights into the fundamental processes that govern auditory perceptual continuity (Warren et al., 1988).

3.2.2. Gap transfer illusion

Continuity illusions have been observed not only in complex sounds like speech, familiar music clips and sound textures (Müller et al., 2013), but also in scenarios involving multiple sound sources, such as rainfall or crowds. These illusions can occur in both temporal and spectral domains (McWalter and McDermott, 2019), as demonstrated by the *gap transfer illusion* (Nakajima et al., 2000), where a temporal gap filled by a brief tone can create the perception of a continuous pitch path. Remarkably, spectral continuity can also be perceived without a bridging tone when

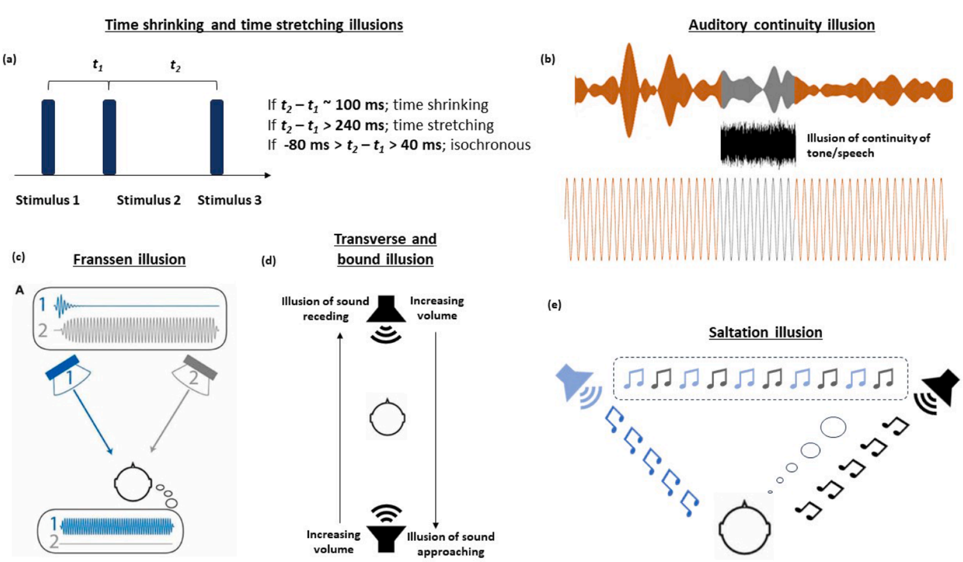


Fig. 2. Examples of temporal and spatial auditory illusions. (a) Time shrinking and time stretching illusions where the difference in the temporal gaps between three tones determine if the second time gap is perceived as shorter, longer or the same duration as the first time gap. (b) Auditory continuity illusions, where speech or pure tones separated by a temporal gap filled with a loud noise is perceived as a continuous stimulus. (c) Franssen illusion, where the onset and steady-state parts of a stimulus are played separately, the participant localises the sound to be coming from the source playing the onset, (d) Transverse and bound illusion where a sound that is increasing or decreasing in volume between two speakers each placed in front or behind a person is perceived to be approaching or receding respectively, (e) Saltation illusion, where different stimuli played through different sources are perceived as interleaved.

two frequency glides with overlapping spectral edges are presented sequentially (Remijn and Nakajima, 2005). Overall, auditory continuity illusions reveal how the auditory system integrates stimuli over time, maintaining perceptual continuity even amid interruptions (Riecke et al., 2008). This phenomenon, commonly experienced in noisy environments like restaurants, underscores the auditory system's resilience to noise and its ability to "fill in" missing sensory information. The core principle of perceptual restoration is that once the context of a sound is established, a stimulus that sufficiently masks a temporal gap or discontinuity and contains relevant spectral information can trigger the "filling-in" process and sustain the continuity illusion (Bregman, 1990; Bregman and Dannenbring, 1977; Warren et al., 1972).

4. Understanding auditory spatial localisation using illusions

Another fundamental function of the auditory system is auditory spatial awareness – the awareness of the presence, distribution and interaction of sound sources in the surrounding space. The entire experience of auditory spatial awareness depends on the physiological state of the listener's hearing abilities, their auditory exposure, knowledge of listening strategies, familiarity with the surrounding environment and degree of involvement in the listening activity. There are primarily three characteristics of auditory spatial awareness (Letowski and Letowski, 2012) (i) auditory localisation – the ability to determine *where* a particular sound is coming from, (ii) auditory distance estimation – the ability to make a judgement on how near or far the source of the sound is from the listener and (iii) spaciousness perception – the perception of the size, type and number of sound sources along with where it's coming from in 3-dimensional space.

4.1. *Understanding the fundamentals of auditory spatial localisation using illusions*

Unlike pitch and rhythm, spatial localisation in not processed by the inner ear itself. It heavily depends on the size and shape of the head and external ears that create a very specific head-related transfer function (HRTF) for each person customising the characteristic directional properties of each person's auditory spatial perceptual experience (Letowski and Letowski, 2012; Møller et al., 1995). The location of the ears on either side of the head creates a difference in timing in which the stimuli arrive at the two ears. This is called intra-aural time difference (ITD) and intra-aural phase differences for sine waves (IPD). A very basic illusion perceived by participants when sound is presented through headphones without pre-processing using the HRTF is called *in-head illusion*, where participants localise the source of the sound as coming from within the head. This is also perceived in sound field in an anechoic chamber when there is a single sound source or multiple sound sources at equal distances. It can also be manipulated to be perceived on one side (Letowski and Letowski, 2012). Although there is not a lot of research on these in-head phantom sounds, future research could delve into whether they be used as another model to study clinically present in-head phantom sounds such as tinnitus.

4.2. *Understanding auditory spatial localisation across different planes using illusions*

4.2.1. *The precedence effect*

One of the ways the auditory system localizes sound sources is to rely on its response to sound cues that arrive first, which typically correspond to the direct path from the sound source to the ears (Brown et al., 2015; Wallach et al., 1949). Similar sounds that follow closely, often arriving milliseconds later as reflections, are generally suppressed by the auditory system in favour of the first-arriving sound. This phenomenon is exemplified by the *precedence effect* or law of the first wavefront (Wallach et al., 1949) – which is a psychoacoustical phenomenon of the brain to make decisions on sound and echo perceptions and locations of sounds coming from a source beyond a certain distance. The precedence effect is particularly important when hearing in enclosed spaces wherein the location of the primary sound may be accurately localised without its interference by the sound's echoes.

4.2.2. *Haas illusion*

The precedence effect is leveraged in a spatial illusion called the *Haas illusion* (Haas, 1951). In this auditory localization illusion, two identical sounds presented in rapid succession (the lead sound followed by the lag sound) are perceived as a single, fused sound, with the perceived source location determined by the lead sound (Wallach et al., 1949). Further research confirmed this effect using a single coherent sound reflection, reinforcing the concept that the auditory system prioritizes direct sound for spatial localization (Haas, 1951, 1972). Psychophysical studies have systematically explored factors influencing the precedence effect, consistently highlighting the dominance of initial direct sounds in determining sound source locations (Brown et al., 2015). The Haas effect is primarily used in music production to create a widened stereo sound out of a mono recording.

4.2.3. *The Franssen illusion*

Another illusion leveraging the precedence effect is the *Franssen illusion* which underscores the auditory system's reliance on onset information for sound localization. This illusion involves two loudspeakers placed in front of a participant, where one (left) speaker only plays the onset of a sound, while the remainder steady-state part of the sound is emitted from the other speaker (right) in such a way that the sum of the two is a sine wave with a regular onset. When the steady-state part is turned off, the left speaker plays a complimentary transient signal such that the sum of the left and right is zero. Remarkably, participants perceive the sound as originating from the left loudspeaker only, despite the majority of the sound coming from the right, demonstrating the overriding influence of onset cues in auditory localization (Franssen, 1962) (Fig. 2c). A second version of the Franssen illusion has the right speaker go on indefinitely, and the participants perceive that the sound on the left speaker is playing indefinitely (Hartmann and Rakerd, 1989). However, it was also found that the Franssen illusion failed in an anechoic chamber (Hartmann and Rakerd, 1989). This was attributed to the "plausibility hypothesis" that participants subconsciously reweighted the steady-state intra-aural cues since they don't know how to compensate for room reflections in localisation of sounds without an onset. In an anechoic chamber with no echoes and room-reflections, participants are able to perfectly locate sounds without an onset showing the importance of room's acoustical design in the localisation of sounds. Furthermore, the Franssen illusion was also shown to fail for noise stimuli showing that the localisation also depended on the spectral density of the signal (Hartmann and Rakerd, 1989).

4.2.4. *Transverse-and-Bounce auditory illusion*

ITDs and ILDs provide left-right localisation cues at low and high frequencies respectively. However, binaural cues are limited in localisation in the vertical and front-back cues inducing illusions in these planes. The *Transverse-and-Bounce auditory illusion*, observed by Bainbridge et al. (2015), exploits the inherent ambiguity in localizing sounds from front versus back, due to similar ITDs and ILDs. In their experiment, a sound with varying volume was played through two speakers positioned in front of and behind participants. When the volume increased from the front to the back, participants perceived the sound as approaching, and when it decreased, the sound was perceived as receding (Fig. 2d). By synchronously varying the volume in both speakers, the researchers created the illusion of four possible movements: front-to-back, back-to-front, front-to-front (bouncing off the participant), or back-to-back (bouncing behind the participant). The participants showed a significant preference for perceiving the sound as moving in a transverse manner—either front-to-back or back-to-front—suggesting a bias towards interpreting continuous trajectories

(Bainbridge et al., 2015). Additionally, participants often alternated between perceiving the sound as moving front-to-back and then back-to-front, which may reflect the auditory system's tendency to maintain a consistent interpretation of sound movement over short periods (Spelke et al., 1995).

4.3. Understanding auditory distance estimation with the auditory distance illusion

In addition to ITD cues, the difference in loudness of the sound that reaches the two ears – called intra-aural level difference (ILD) – plays a crucial role in estimating the distance to a sound source. This ability is further mobilised by the *auditory distance illusion*, as demonstrated by Butler and Naunton (1964). Their study revealed that the perceived auditory distance can be manipulated through the interplay between loudness cues and the brain's interpretation of auditory space (Butler and Naunton, 1964). Generally, louder sounds are perceived as closer; however, this perception can be altered by the presence of other sounds. By varying the loudness relationship between sounds emitted from a loudspeaker and those delivered through headphones, Butler and Naunton induced an illusory shift in perceived distance. Specifically, an increase in overall loudness in the auditory environment caused the perceived location of the sound source to move farther away, and vice versa. This auditory distance illusion highlights the reliance of the auditory system on relative loudness cues to shape our perception of sound source distance.

4.4. Understanding auditory spaciousness perception using illusions

4.4.1. Auditory saltation illusion

Other aspects of spatial perception such as localisation of multiple sound sources is also studied through an illusion exposing the constraints of auditory spatial perception. The *auditory saltation illusion*, first observed by Bremer et al. (1977), illustrates the reduced precision of spatial localization for sounds presented in rapid temporal succession (Bremer et al., 1977). In this illusion, when brief sounds are presented in rapid succession from two distinct locations, listeners perceive the sounds as jumping across intervening spaces (Fig. 2e). For example, when a series of clicks lateralized to the left by an ITD is followed by another series of clicks lateralized to the right, participants perceive them as originating from a smoothly varying series of locations spanning the true source locations (Bremer et al., 1977; Hari, 1995; Shore et al., 1998). The distinct sources of the two stimuli are perceived as fused, with the sound seeming to move between anchor points and emit clicks periodically as it does so (Hari, 1995). Research has shown that this illusion is more pronounced when the sounds are difficult to individualize and localize, and it is highly dependent on the inter-stimulus interval (ISI) between the sounds. Consistent evidence indicates that the auditory saltation illusion is strongest when sounds are presented with an ISI of $<$200 ms (Hari, 1995; Phillips and Hall, 2001; Shore et al., 1998). This phenomenon highlights the temporal constraints of the auditory system in accurately localising sounds, particularly when dealing with rapidly sequential stimuli from different locations, leading to this illusory spatial perception.

4.4.2. Solitaire illusion

Understanding spatial perception with respect to the number of sound sources and type of sounds can be studied using the *solitaire illusion.* The solitaire illusion is a visual illusion where the arrangement of elements influences perceived numerosity (Frith and Frit, 1972). In the auditory domain, Prpic and Luccio (2016) reported a similar illusory perception. They created audio sequences with piano and drum sounds, manipulating the pitch arrangement of the piano tones. Participants perceived more drum sounds when piano tones were perceived as separated by drum sounds, suggesting they were grouped into smaller units, whereas they perceived more piano sounds when piano tones were perceived as a single group (Prpic and Luccio, 2016). The principles of contiguity and spatial separation, first proposed by Wertheimer, are thought to explain the Solitaire Illusion. According to the hypothesis, elements that are spatially clustered or contiguous are perceived as a single large unit, whereas those that are separated are individual small units (Wertheimer, 1923). The number of elements in several small units is underestimated when compared to a single large unit of elements. This difference in grouping explains why more drum sounds were perceived when they were arranged as a single unit separating the piano tones, which were perceived as several small units.

5. Understanding interactions between spectral, temporal and spatial domains using illusions

Spectral, temporal and spatial processing in the auditory system is not always independent. The different components interact with one another and influence the perception in these domains. The following illusions illustrate this interaction resulting in a percept combining the different domains as seen in the different types of auditory movement-based illusions. Movement-based illusions use ITD and ILD cues together with other spectral cues to perceive the illusion of continuous movement of sound across space that is segmented by an array of segregated speakers.

5.1. Octave illusion

A well-studied auditory illusion induced by harmonic complex tones is the *octave illusion.* Although the illusion is itself may be considered spectral, it is induced by combining binaural perceptual cues. First reported by Deutsch in 1974, the octave illusion was induced when two tones separated by an octave, such as 400 Hz and 800 Hz, were presented dichotically and simultaneously through stereo headphones in a repeating pattern (i.e., right ear receives high tone while left ear receives and then switches) with each trial lasting 250 ms (Deutsch, 1974). This pattern resulted in the perception of a single tone that alternated between ears but also a fixed localisation of the high pitch even when the headphones were reversed (Deutsch, 1974).

The octave illusion reveals important insights into the complexities of auditory perception, particularly the interplay between pitch perception ("what") and sound localisation ("where"): The perceived pitch, "what", is determined by the dominant ear, whereas sound source, "where", is located in the ear receiving a high tone (Chambers et al., 2004; Deutsch and Roll, 1976). While the octave illusion highlights the intricate nature of how the auditory system determines the location of a sound source, there is evidence that shows that individuals localised the illusion toward the lower frequency component (Chambers et al., 2002). In addition, right-handed individuals showed a strong tendency to localize the higher tone in the right ear, while left-handed individuals did not demonstrate this preference (Deutsch, 1974), suggesting a relationship between the perceived location of the tones and hemispheric dominance that is often linked to handedness. Thus, these inconsistent findings prompt further investigation into the interplay of different localization cues and a more unified model.

5.2. Auditory motion illusion

One of the well-studied aspects of auditory illusions in the spatial domain is the perception of speed and movement. Examples of these would be the *auditory motion illusion, auditory flash-lag illusion* and the infamous *Doppler Illusion. Auditory motion illusion* involves the perception of speed, manipulated through amplitude modulation (AM) frequency. Senna et al. (2017) demonstrated that AM frequency serves as a cue for perceiving auditory speed. In their study, moving rattling sounds were generated by applying AM to pink-noise carrier signals and played through a circular array of eight loudspeakers, creating the perception of circular motion. Sounds with higher AM frequencies were perceived as moving faster than those with lower AM frequencies. Furthermore, higher AM frequencies induced a stronger motion aftereffect, where a subsequent stationary sound appeared to move in the opposite direction. These findings suggest that the auditory system uses AM frequency as a relative cue for interpreting perceived auditory motion, with differences in AM frequency directly correlating with perceived speed (Senna et al., 2017). This aligns with the concept that the perceptual system exploits natural scene statistics to infer the state of the world (Parise et al., 2014; Senna et al., 2015).

5.3. Auditory flash-lag illusion

The "flash-lag" illusion was originally studied in the visual domain where a moving object and a stationary object that flashes in the same location are perceived as misaligned (Nijhawan, 1994). In the auditory "flash-lag" illusion, a 1 second tone was swept over a two-octave frequency range (from 1 to 4 kHz) in a log-linear frequency glide in one ear. In the middle of this sweep, a brief 40 ms tone burst was played to the other ear. This tone burst was initially 2 kHz, and the task was to judge if this tone was higher or lower than the sweeping frequency. An adaptive staircase using Quest is used to adjust to the point where the swept frequency and the tone burst are perceived to be at the same frequency. However, the perception of equality of the sweeping frequency and the tone burst does not occur at the physical equality of the frequency characteristics of the two stimuli thereby creating the illusion (Alais and Burr, 2003; Arrighi et al., 2005). This auditory illusion can also be induced by a translating sound across space and a static tone. Hubbard (2014) categorized potential factors influencing the visual "flash-lag" illusion into two main categories: characteristics of the stimuli (e.g., speed, duration, distance, direction, luminance, and contrast) and characteristics of the observers (e.g., attention, movement, perceptual organization, perceptual set, and conceptual knowledge) (Hubbard, 2014). However, research into the characteristics of the auditory "flash-lag" illusion remains limited and thereby the understanding of how it may be applied to understand auditory perception.

5.4. The Doppler effect

In 1996, Neuhoff and McBeath observed the *Doppler illusion* – the perception of a rising pitch as a sound source approaches, even though the frequency actually decreases (Neuhoff and McBeath, 1996). This phenomenon stems from the Doppler effect, which is the change in perception of the frequency of a wave (increase or decrease) that arises when there is relative motion between a wave source and an observer (approaching or leaving). This effect is more pronounced in dynamic tones compared to static tones. Notably, while complex sounds could lead to a stronger illusory perception compared to the pure tones, the intensity of the stimulus also has been shown to produce a pronounced effect for pure tones but less pronounced for complex sounds (Mcbeath and Neuhoff, 2002). In addition, the Doppler effect is not uniform across the entire auditory frequency range. The effect is stronger at higher frequencies and for more complex waveforms (Väljamäe, 2009; Vesely et al., 2018). Overall, the Doppler illusion indicates that the perception of pitch is not solely determined by the frequency of a sound but can be influenced by the relative motion between sound source and listeners.

6. Understanding higher-order auditory and multi-sensory processing through illusions

Higher order auditory processing involves not only the comprehension of hearing and speech but also the interaction with other sensory modalities. For example, understanding speech and more particularly conversation is not just a purely auditory phenomenon, but also combines visual cues about lip movements, facial expressions and other non-verbal cues to produce a continuous and information-rich perceptual experience. By the virtue of integrating inputs from different sensory modalities, the perception of one sensory modality can be altered or misinterpreted due to the influence of another which can help understand the mechanism behind different higher order processes. The illusions explained below demonstrate this phenomenon:

6.1. Sound-induced flash illusion

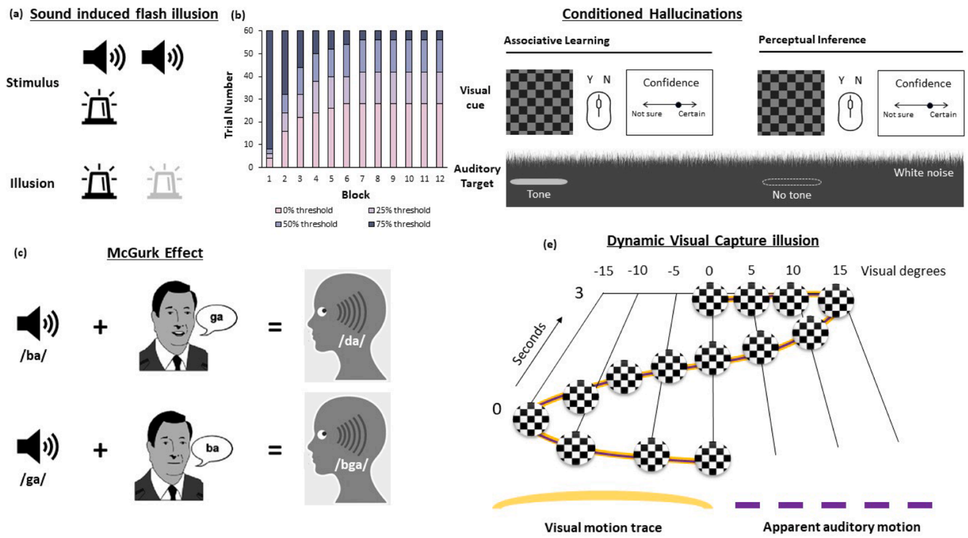
The interaction between the auditory and visual modalities is a long-researched topic. Simple illusions like the *sound-induced flash illusion (SIFI)* and *conditioned hallucination* bring out the multisensory interaction between audio-visual domains showing how one modality can trigger illusion in another (Hirst et al., 2020). It also showcases the ability of the brain to integrate multisensory objects into a model which allow for prediction of the occurrence of the object itself. When one of the components is not presented, the brain computes the optimal probabilistic occurrence of the multisensory object itself and induces an illusion of one or the other. The SIFI entails the rapid presentation of a single visual stimulus, a “flash” with two auditory stimuli, “beeps” resulting in the perception of two flashes (Hirst et al., 2020) (Fig. 3a). Several theories have supported the generation of this illusion including the directed attention hypothesis that proposes that the attended modality drives the perception, the modality-appropriateness hypothesis suggesting that the sense most appropriate for the task drives the perception and maximum likelihood estimation model that proposes that the sensory information is weighted according to its relative reliability and precision (Hirst et al., 2020).

6.2. Conditioned hallucination

In the *conditioned hallucination* illusion, a suprathreshold auditory target embedded in a noise floor is presented together with a visual cue (Powers et al., 2017). This presentation is repeated through a non-linear process where, through the course of the experiment, the proportion of subthreshold and absent target trials were increased. It was observed that when only the visual cue was presented, an illusory experience of the missing target tone was perceived (Fig. 3b). This is based on the idea that a pavlova-type conditioning of perceiving a tone when a visual cue is presented is built through the course of the experiment, and when only the cure and not the target is presented, it is expected or predicted to be present, and the brain perceives the missing stimulus as a conditioned response (Powers et al., 2017). It is also shown that people who tend to build stronger associative models of audio-visual integration are more likely to perceive the illusion (Kafadar et al., 2022). This illusion has been used as a model to study its related hypothesis in different clinical populations, particularly with auditory hallucinations (Yasoda-Mohan et al., 2024).

6.3. The McGurk effect

One of the most notable illusions that arises from multisensory integration in the speech domain is the *McGurk effect*. McGurk and MacDonald (1976) observed that a visual cue of /ga/ paired with the auditory stimulus of /ba/ could lead to the illusory perception of /da/, whereas the responses to the opposite pairing of visual /ba/ with auditory /ga/ contained combination responses such as /bga/ (McGurk and MacDonald, 1976) (Fig. 3c). Even though there is a physical discrepancy of /g/ being a velar consonant with mouth open and /b/ being a bilabial consonant with mouth closed, the participants do not detect intermodal conflict but experience a phoneme that does not represent the played auditory signal. McGurk and MacDonald (1976) argued that: “*… In a ba-voice/ga-lips presentation there is visual information for [ga] and [da] and auditory information with features common to [da] and [ba]. By responding to the common information in both modalities, a subject would arrive at the unifying percept [da]…*”. This is consistent with the finding that task-relevant visual information could enhance the auditory perceptual performance (Møller et al., 2018; Su, 2014).



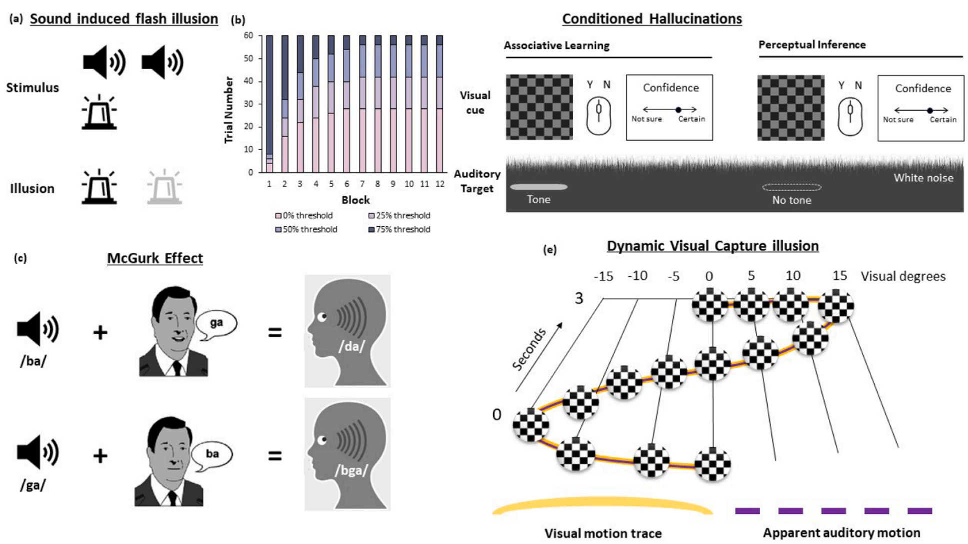


Fig. 3. Examples of multisensory illusions. (a) Sound induced flash illusion, where a multisensory stimulus (sound + flash) followed by a single sound induces the illusion of a flash, (b) Conditioned Hallucinations, where the association of a visual cue and auditory target produces the illusion of the target when cued only by the visual stimulus with an absent auditory target, (c) McGurk Effect, where the auditory input of /ba/ combined with the visual cue of /ga/ produces the percept of /da/ , however when combined with the visual cue of /ba/ produces the percept of /bga/, (d) dynamic visual capture illusion, where the visual trajectory of a stimulus combined with an auditory stimulus gives the illusion of the movement of the sound.

Essentially, the McGurk effect demonstrates that speech perception is not solely driven by acoustic signals; instead, it's a multimodal process where visual cues from a speaker's mouth movements contribute significantly to our understanding of spoken language, particularly in noisy environments where the acoustic signal might be ambiguous (Getz and Toscano, 2021). Therefore, it highlights the influence of higher-level cognitive processes on auditory perception.

6.4. Speech illusion with concurrent texts

The *speech illusion with concurrent texts* is another auditory illusion that is largely rooted in the brain's multisensory integration capabilities. This auditory illusion, first observed by Frost et al. (1988), arises when a spoken word is masked by noise having the same amplitude envelope, but a concurrent printed version can enhance the clarity (Frost et al., 1988). Following studies reported the similar illusory phenomenon by using the corresponding video records of a speaker's mouth (Repp et al., 1992). The audiovisual integration can occur even if the auditory and visual signals are not perfectly synchronized, as long as the temporal shift is within a certain threshold (e.g., 50–100 ms) (Repp et al., 1992). The generation of the speech illusion with concurrent texts heavily relies on the contextual association or congruency and is automatically processed (Frankish, 2008; Frost et al., 1988), which can be supported by the reduced illusory perception if used auditory inputs with no envelope information or visual nonwords (Frost et al., 1990). A version of this is that even when the same auditory stream is presented with different visual textual sentences, the observer hears the vocalisation of the visual text. Analogous to the McGurk effect, this auditory illusion highlights that the auditory perception, particularly of speech, is largely shaped by top-down influences, such as lexical knowledge and phonetic encoding, which interact with bottom-up acoustic information to generate the final percept. Additionally, visual information also biases the perceptual auditory experience to expect to hear what they see.

6.5. The ventriloquist's illusion

Audio-visual disintegration could also arise where the perceived location of a sound is misattributed to a visible but soundless source, as known as the *ventriloquist's illusion* (Bertelson and Radeau, 1981; Driver, 1996; Radeau and Bertelson, 1974; Thurlow and Jack, 1973). This illusion exemplifies how the brain integrates the conflicting information between different sensory modalities by prioritizing the more reliable or dominant ones, in which case the visual over auditory (Bruns, 2019). Ventriloquism exploits this by synchronizing the movements of a puppet's mouth with the speech of the ventriloquist, leading the audience to perceive the voice as being produced from the puppet rather than the ventriloquist. The ventriloquist's illusion has been experimentally explored. Specifically, participants are exposed with a synchronous but spatially discrepant audiovisual stimulus and are required to localise the source of auditory cues. The ventriloquist's illusion occurs when they perceive the auditory stimulus closer to the visual stimulus rather than the actual source (Bertelson and Radeau, 1981). Following studies further investigating the psychoacoustic characteristics of the ventriloquist's illusion, can be broadly categorized into contextual, spectral, spatial and temporal elements.

6.6. Dynamic visual capture illusion

The *dynamic visual capture illusion* is an auditory illusion where an auditory stimulus is perceived to be moving in the same direction as a visual stimulus, even though the auditory stimulus is static or even moving in the opposite direction (Mateeff et al., 1985). Such an illusory movement of auditory source was first observed by Mateeff et al. (1985) and validated by following studies (Alink et al., 2008; Kitajima and Yamashita, 1999; Soto-Faraco et al., 2002; Vroomen and de Gelder, 2003). The velocity of sound movement is positively associated with the sense of speed of the visual movement (Kitajima and Yamashita, 1999). Thus, this visual capture effect on auditory motion might be a strong indicator of direct cross modal interactions in motion perception (Hidaka et al., 2015) (Fig. 3d). Taken together, compared to other spatial auditory illusions, these two illusory models probe the auditory spatial perception as an active process that incorporates visual information for localisation, highlighting the brain's tendency to integrate information from different senses to form a coherent perception of the environment.

6.7. Speech-to-song illusion

Deutsch et al. (2011) reported an intriguing auditory illusion where a repetitive speech segment is able to induce a perceptive transformation from speech to song. They created this *speech-to-song illusion* without altering any signals or adding any musical components but simply by repeating the speech phrase (Deutsch et al., 2011). Notably, the speech-to-song illusion was observed not only in other languages such as German, Mandarin and Thai (Falk and Rathcke, 2010; Jaisin et al., 2016) but also in other nonspeech sounds, such as sequences of tones (Margulis and Simchy-Gross, 2016) or environmental sounds (Simchy-Gross and Margulis, 2018), confirming the robustness of this auditory illusion. The underlying mechanism of the speech-to-song illusion may lie on the distinct perceptual processing of speech and music. Exact repetition of stimuli may cause the disinhibition of the pitch processing during the perception, which renders the pitch more salient that sounds like a song (Deutsch, 2019).

Previous studies have investigated the psychoacoustic characteristics of the speech-to-song illusion, including the linguistic factors, acoustic factors and individual differences. For the linguistic factors, individuals with tonal native language (e.g., Mandarin, Thai) experience a significantly weaker speech-to-song illusion compared to those with non-tonal native languages (e.g., English, German) (Jaisin et al., 2016). In addition, the pronunciation difficulty of the speech phrase relative to their native language is positively linked with the strength or susceptibility of the illusory perception, which might result from the decreased capability to capture acoustically complex languages (Margulis et al., 2015). Similarly, the speech-to-song illusion is stronger if the participant does not understand the language of the speech phrase (Jaisin et al., 2016). For the acoustic factors, syllables with greater fundamental frequency stability and favouring a metrical interpretation are associated with speech samples that tend to elicit the speech-to-song illusion, though it is not a primary driver and does not guarantee the illusion (Tierney et al., 2013). Alternatively, one study also demonstrated that individuals who are more sensitive to rhythmic cues are more likely to experience the speech-to-song illusion (Falk et al., 2014). However, the strength of the speech-to-song illusion is not affected by music training (Falk et al., 2014) or aging (Mullin et al., 2021).

Table 1 provides a summary of all auditory illusions described in this manuscript.

7. Discussion

Illusions have been historically used to understand the functioning of sensory domains because of their ability to resolve conflicts when the brain receives contradictory sensory information. Illusions are usually experimentally induced however, our everyday perception itself may be argued to be an illusion (Tiehen, 2023). In the visual domain, it is well known that although the blind spot in the retina receives no input creating a sensory discrepancy to the input to the brain, this blind spot is not perceived in the visual field (Spillmann et al., 2006). Another example is the perception of a motion picture – where individual categorical images when rolled at 24 frames per second, it is perceived as continuous motion (Anderson and Anderson, 1980). Similar examples are also showcased in the auditory domain with the perception of speech. The categorical changes in different phonemes although extremely critical are not noticed leaving a seamless perception of a person's speech (Massaro and Cohen, 1983).

In the current manuscript, we build on the work done by Scharine and Letowski in using auditory illusions to understand different psychophysical characteristics of the auditory system in the spectral, temporal, spatial and multisensory domains (Scharine and Letowski, 2009). Although illusions have been studied independently, bringing them together under the umbrella of auditory processing and their related psychoacoustics has broadly three reasons – (i) using them as proxies or causal human models for understanding neuropathology with and without a psychiatric component, (ii) using them as causal models to understand the mechanisms of aesthetic perception (iii) leveraging the working of the auditory system to enhance perception of sounds, either in everyday life or in an aesthetic context and (iii) expanding perception beyond the realm of conventional human perception using virtual reality.

7.1. Use of auditory illusions to understand mechanisms of clinical pathologies

The use of auditory illusions to study neurological disorders has been implemented in several occasions. It was shown that perceiving the missing fundamental frequency was impaired in patients with Alzheimer's disease (AD) (Abe et al., 2018). The accuracy on the task correlated with changed in grey matter volume in the insula, temporal pole, inferior frontal gyrus and entorhinal cortex, some of which are known to be impacted in AD. Similarly Shepards Tones have been utilised in understanding the development of cognitive skills in children where authors found that receptive language and abstract reasoning increased with age and correlated with the strength of the Shepards tones perception, however not more than the effect of maturing auditory processes (Chouinard et al., 2020). Similarly, different types of auditory temporal domain illusions and effect of ordering of auditory objects may be used in the study of mechanism involved in central auditory processing disorders (Chowsilpa et al., 2021) and autism spectrum disorders (Kwakye et al., 2011) whereas auditory spatial domain illusions may be employed to test the effect of age and hearing loss on auditory processing (Roberts Richard and Lister Jennifer, 2004).

Furthermore, the Zwicker Tone illusion, conditioned hallucination and McGurk effect, to name a few, have been extensively studied under

Table 1
Summary of auditory illusions across spectral, temporal spatial and multisensory domains.

Auditory domain	Auditory sub-domain	Name of illusion	Explanation	Application
Spectral	Tonotopic pitch processing	Zwicker tone illusion	Auditory after-image of missing central frequency following continuous exposure of broadband noise with a spectral notch	Postulated as a human model for tinnitus – continuous ringing in the ear
	Processing sound harmonics	Periodicity pitch or missing fundamental	Perception of fundamental frequency even when it is deleted from a complex harmonic tone	Seamless speech perception through analogue telephones lines which cuts off fundamental frequency of human voice
		Circular pitch illusion or Shepards Tone	12-tone sequence spanning one octave separate by a semitone presented with a progressively decreasing/increasing volume producing an illusion of one pitch that is continuously increasing/decreasing	Understanding adaptive mechanisms of the brain that is involved in perception of auditory features
Temporal	Temporal discrimination and ordering	Time shrinking illusion	Short time gaps (< 200 ms) preceded by shorter time gaps is perceived shorter than original	Music production particularly in the digital space
		Time stretching illusion	When difference between the two gaps > 240 ms, the second gap is perceived to be longer than original	
		Isochronicity illusion	When the difference between two gaps is between −80 ms to 40 ms the two gaps are perceived as same length even when it is not. This can also happen when sounds are presented within and outside a particular context	
	Temporal masking and Temporal integration	Auditory continuity illusion	A sound with a temporal gap when played with a loud masking noise is perceived as a continuous tone	Example of how the brain "fills-in" for missing information
		Gap transfer illusion	A temporal gap filled by a brief tone creates the perception of a continuous pitch	
Spatial	Sound localisation	In-head illusion	Perception of sound inside the head when sound play is customised with a head-related transfer function	Applied in customising sound play in AirPods to a person's head size
		Precedence Effect	The echoes of a sound is automatically reduced by the ears and only the dominant sound is perceived	These are particularly important phenomenon when designing acoustic environments such as recording studios, movie theatres, concert halls etc.
		Haas illusion	Based on the precedence effect – rapid sounds in succession is perceived as one	
		Franssen illusion	When onset and steady state portions of a tone are separated and played through different speakers, participants localise the sound to that playing the onset	
		Transverse-and-Bound illusion	Varying volume between two speakers placed in front and behind a participant produces an illusion of sound approaching and receding when played from the front and back speakers respectively	
		Octave illusion	Two tones separated by an octave presented dichotically in a repetitive pattern of high and low in either ears results in the perception of a single tone that alternates in ears and pitch	
	Distance estimation	Distance illusion	The illusion that louder sounds feel closer and softer sounds feel farther away	
	Spaciousness	Auditory Saltation illusion	The perception of reduced spatial localisation when tones are presented from different sources but in rapid succession	The experience of sitting in a crowded café or bar and unable to place where someone else is talking from
		Solitaire illusion	When different types of sounds are played together, manipulating the pitch of one groups their perception in into sub units and separates them from other sounds	Digital music production where instrument levels have to be adjusted to produce a coherent perception of song.
Interaction between spectral and temporal	Movement	Auditory motion illusion	Amplitude modulated pink noise when played through a circular array of eight speakers creates the illusion of circular movement	These types of illusions are used to produce special effects in movies and animations. The Doppler effect has been adapted in ultrasound to measure the distance of moving objects.
		Auditory flash-lag illusion	A one second tone swept over two octaves in a log-linear frequency glide is played in one ear and a short burst is played to the other ear. When this set up is modulated to until when the perception of pitch of the tone and the burst are equal, when in real they are not	
		The Doppler Effect	When the frequency of a tone is increased or decreased across an array of speakers, this creates an illusion of a moving sound	
Multi-sensory processing		Sound-induced flash illusion	When two tones are played synchronously with one flash, the participant perceives two flashes	These illusions may be used as causal models to understand multisensory integration in different populations – both healthy and patient groups.
		Conditioned Hallucination	When a tone is played together with a visual cue repetitively, and when the cue is played by itself without the tone, an illusion of the tone is perceived	
		The McGurk Effect	A visual cue of a talker saying /ga/ together with an auditory stimulus of /ba/ produces a percept of /da/	
		Speech illusion with concurrent texts	The same auditory stimulus paired with a noise masker cued with different visual texts, produces a percept of the different texts	

(*continued on next page*)

Table 1 (*continued*)

Auditory domain	Auditory sub-domain	Name of illusion	Explanation	Application
		Dynamic Visual capture	An auditory stimulus is perceived to move in the same direction as the visual stimulus even though the auditory stimulus is static.	
		Ventriloquist's illusion	An audio-visual disintegration illusion where a puppet's mouth movements combined with the auditory stimulus creates an illusion of the puppet talking rather than the artist.	Famously used in puppetry
		Speech-to-song illusion	Repetitive categorical speech stimuli are perceived as a song	Maybe used to understand the basis of aesthetic perception in the brain

different pathological contexts. The Zwicker tone illusion has been studied as a proxy for tinnitus, the continuous phantom auditory perception given its psychoacoustic characteristics (Leske et al., 2014; Mohan et al., 2020; Norena et al., 2000). Similarly the conditioned hallucination paradigm has been applied to understand the mechanism of action of generating hallucinations both in a healthy adult group and in a group with schizophrenia (Kafadar et al., 2022; Powers et al., 2017). Interestingly, it has also been used to disentangle the mechanism of hallucinations in people with and without psychosis (Powers et al., 2017). The McGurk Effect has been used to study audio-visual speech integration in children with learning disabilities (Boliek et al., 2010), adolescents and adults with schizophrenia (Pearl et al., 2009; White et al., 2014) and bipolar disorder (Yordamlı and Erdener, 2018). Similarly, the Ventriloquist effect has been used to understand if the absence of visual input due to hemispatial neglect can still direct people's perception towards the illusory location of the auditory stimulus giving us a deeper understanding of audio-visual integration (Bertelson et al., 2000; Calamaro et al., 1995). Multisensory integration has also been studied using the sound induced flash illusion in both patients with multiple sclerosis (Yalachkov et al., 2019) and healthy adults induced with parietal lobe disruption through TMS (Kamke et al., 2012) as a lesion deficit model. Finally, increased sensitivity to movement in the flash-lag illusion has been associated with symptom severity in patients with schizophrenia (Zeljic et al., 2021), although it is debated (Kaliuzhna et al., 2020). In this way other illusions could also be used as either a human model or as a causal tool to understanding the mechanisms of a disorder. Furthermore, the Doppler Effect has been studied to trigger an instantaneous escape mechanism in bats (Yoshida et al., 2024) in addition to being used in radar, particularly in the ultrasound frequencies.

7.2. Use of auditory illusions to understand the mechanisms of aesthetic perception

From an evolutionary perspective, music is viewed as a by-product of natural selection in cognitive domains such as language, auditory scene analysis, habitat selection, emotion and motor control (Hauser and McDermott, 2003; Pinker, 2005). The importance of aesthetics in perception is an increasingly studied topic, especially from a neuro-aesthetic perspective. Particularly, while scanning the external world, aesthetics adds an important dimension of valence which is foundational to how we evaluate objects (Hellström, 2024). In this regard, an aesthetic sense to perception is key to assigning value, salience and determining stimulus selection. This is demonstrated by Zaidel who stated, *"The human-specific aesthetic response itself likely rests on evolutionary modification of the biologically tuned attentional system"* (Zaidel, 2020). From this angle, illusions like the speech-to-song gives us insight on the importance of aesthetics in making sense of the world. This has also been demonstrated in both treatment and educational spaces – music therapy has been used as a treatment for stuttering (O'Donoghue et al., 2020), Parkinson's Disease (Machado Sotomayor et al., 2021), aphasia (Schlaug et al., 2010) and autism spectrum disorders (Marquez-Garcia et al., 2022). Furthermore, arts-based practices are now increasingly used in developing pedagogies in basic (Ward, 2013) and higher education streams (Rieger et al., 2016). Therefore, illusions of an aesthetic nature can give us a causal tool to delve into the neural mechanisms underlying aesthetic perception.

7.3. Use of auditory illusion in entertainment – design of sound and acoustic environments

A third way we might leverage the power of auditory illusions, as we have, is in the entertainment business. Different illusions like the circular pitch illusion (Shepard, 1964), time-stretching, time-shrinking and isochronous illusions (Nakajima et al., 2004; Sasaki et al., 2010) are crucial concepts to keep in mind while producing music, particularly in the digital space. Taken together with the Franssen effect (Hartmann and Rakerd, 1989) and Haas illusion (Haas, 1972) which are variations of the precedence effect, they determine how sound travels in different spaces. Particularly, the Franssen effect is highly dependent on the auditory steady state responses of the room and therefore become a crucial perceptual point to be considered during music production and acoustic environmental design. Similarly, other illusions such as the Transverse-and-Bounce (Bainbridge et al., 2015) which impacts the perception of front vs back, the auditory Saltation illusion (Kidd and Hogben, 2004) which produces a fused percept of different moving components, the Solitaire illusion (Prpic and Luccio, 2016) that biases the perception of one type of source vs the other and the different auditory movement-related illusions that utilise a combination of spectral and temporal components to create a spatial illusion are important variables to be considered when designing a surround sound design. Furthermore, the ventriloquist illusion (Bruns, 2019) as the name suggests, relies on our ability to fuse the lip movement from puppet with the auditory information skilfully delivered by the performer with minimal lip movement and attribute it to the puppet rather than the speaker. This is similar to the famous rubber-hand illusion (Kammers et al., 2009) where ownership of one's own extremity is transferred to a rubber-hand through the integration of visual and somatosensory information.

7.4. Use of illusions in expanding human potential of perception and futuristic technology

Finally, auditory illusions may be a key to expanding our perceptual experience beyond everyday level. In today's fast growing technological world, virtual reality is gaining serious traction. From being used for entertainment and recreation to being used in treatment (Kwon et al., 2023) and educational spaces, virtual reality is starting to become a tool to unlocking new realms and levels of human perception. In the multi-sensory reality of the virtual world, auditory experiences play a crucial role in integration of our sense of self into a human-made environment. From the above examples we understand the importance and our reliance on auditory illusions to create our perception of our auditory environment. Therefore, understanding the psychoacoustic properties of auditory illusions and how they impact spectral, temporal and spatial auditory processing is crucial to creating a virtual world that is closest to the one we live in.

To summarise, auditory illusions offer us an avenue to understand the psychological and neurophysiological implications of manipulating fundamental aspects of the auditory system. It provides us with a direct tool to causally investigate neural mechanisms of health and disease while also enabling us to technologically advance our pursuit of enhancing our perception educationally and recreationally while also pushing boundaries of pre-conceived notions of human perception.

8. Conclusion

The current manuscript therefore explores different domains of auditory processing – spectral, temporal, spatial and multisensory – through the lens of auditory illusions. We go through the experimental and behavioural characteristics of these illusions and their application in different domains. Particularly we stress their importance as causal human models to understand different pathologies. Finally, we also provide other applications to auditory illusions. This way we encourage future research to invest further in this field.

CRediT authorship contribution statement

Anusha Yasoda-Mohan: Writing – review & editing, Writing – original draft, Visualization, Conceptualization. **Feifan Chen:** Writing – review & editing, Writing – original draft, Conceptualization. **Sven Vanneste:** Writing – review & editing, Writing – original draft, Conceptualization.

Data availability

No data was used for the research described in the article.

References

Abe, M., Tabei, K.-i., Satoh, M., Fukuda, M., Daikuhara, H., Shiga, M., Kida, H., Tomimoto, H., 2018. Impairment of the missing fundamental phenomenon in individuals with Alzheimer's disease: a neuropsychological and voxel-based morphometric study. Dement. Geriatr. Cogn. Dis. Extra 8 (1), 23–32. https://doi.org/10.1159/000486331.
Alais, D., Burr, D., 2003. The "flash-lag" effect occurs in audition and cross-modally. Curr. Biol. 13 (1), 59–63. https://doi.org/10.1016/S0960-9822(02)01402-1.
Alink, A., Singer, W., Muckli, L., 2008. Capture of auditory motion by vision is represented by an activation shift from auditory to visual motion cortex. J. Neurosci. 28 (11), 2690–2697. https://doi.org/10.1523/JNEUROSCI.2980-07.2008.
Anderson, J., Anderson, B., 1980. Motion perception in motion pictures. In: de Lauretis, T., Heath, S. (Eds.), The Cinematic Apparatus. Palgrave Macmillan, UK, pp. 76–95. https://doi.org/10.1007/978-1-349-16401-1_7.
Arrighi, R., Alais, D., Burr, D., 2005. Neural latencies do not explain the auditory and audio-visual flash-lag effect. Vision Res. 45 (23), 2917–2925. https://doi.org/10.1016/j.visres.2004.09.020.
Artieda, J., Pastor, M.A., Lacruz, F., Obeso, J.A., 1992. Temporal discrimination is abnormal in Parkinson's disease. Brain 115 (Pt 1), 199–210. https://doi.org/10.1093/brain/115.1.199.
Baguley, D., McFerran, D., Hall, D., 2013. Tinnitus. Lancet 382 (9904), 1600–1607. https://doi.org/10.1016/S0140-6736(13)60142-7.
Bainbridge, C.M., Bainbridge, W.A., Oliva, A., 2015. Quadri-stability of a spatially ambiguous auditory illusion. Front. Hum. Neurosci. 8. https://doi.org/10.3389/fnhum.2014.01060.
Bertelson, P., Pavani, F., Ladavas, E., Vroomen, J., de Gelder, B., 2000. Ventriloquism in patients with unilateral visual neglect. Neuropsychologia 38 (12), 1634–1642. https://doi.org/10.1016/S0028-3932(00)00067-1.
Bertelson, P., Radeau, M., 1981. Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. Percept. Psychophys. 29 (6), 578–584. https://doi.org/10.3758/BF03207374.
Bizley, J.K., Cohen, Y.E., 2013. The what, where and how of auditory-object perception. Nat. Rev. Neurosci. 14 (10), 693–707. https://doi.org/10.1038/nrn3565.
Boliek, C., Keintz, C., Norrix, L., Obrzut, J., 2010. Auditory-visual perception of speech in children with learning disabilities: the McGurk effect. Canadian J. Speech-Lang. Pathol. Audiol. 34 (2).
Bregman, A.S., 1990. Auditory Scene Analysis: The Perceptual Organization of Sound. The MIT Press.
Bregman, A.S., Dannenbring, G.L., 1977. Auditory continuity and amplitude edges. Can. J. Psychol. 31 (3), 151–159. https://doi.org/10.1037/h0081658.
Bremer, C.D., Pittenger, J.B., Warren, R., Jenkins, J.J., 1977. An illusion of auditory saltation similar to the cutaneous "rabbit". Am. J. Psychol. 90 (4), 645–654.
Brown, A.D., Stecker, G.C., Tollin, D.J., 2015. The precedence effect in sound localization. J. Assoc. Res. Otolaryngol. 16 (1), 1–28. https://doi.org/10.1007/s10162-014-0496-2.
Bruns, P., 2019. The ventriloquist illusion as a tool to study multisensory processing: an update. Front. Integr. Neurosci. 13. https://doi.org/10.3389/fnint.2019.00051.
Butler, R.A., Naunton, R.F., 1964. An auditory illusion of distance. Nature 203, 1201. https://doi.org/10.1038/2031201a0.
Calamaro, N., Soroker, N., Myslobodsky, M.S., 1995. False recovery from auditory hemineglect produced by source misattribution of auditory stimuli (the ventriloquist effect). Restor. Neurol. Neurosci. 7, 151–156. https://doi.org/10.3233/RNN-1994-7304.
Carroll, C.A., Boggs, J., O'Donnell, B.F., Shekhar, A., Hetrick, W.P., 2008. Temporal processing dysfunction in schizophrenia. Brain Cogn. 67 (2), 150–161. https://doi.org/10.1016/j.bandc.2007.12.005.
Casini, L., Pech-Georgel, C., Ziegler, J.C., 2018. It's about time: revisiting temporal processing deficits in dyslexia. Dev. Sci. 21 (2). https://doi.org/10.1111/desc.12530.
Chambers, C., Akram, S., Adam, V., Pelofi, C., Sahani, M., Shamma, S., Pressnitzer, D., 2017. Prior context in audition informs binding and shapes simple features. Nat. Commun. 8 (1), 15027. https://doi.org/10.1038/ncomms15027.
Chambers, C., Pressnitzer, D., 2014. Perceptual hysteresis in the judgment of auditory pitch shift. Attention, Percept., Psychophys. 76 (5), 1271–1279. https://doi.org/10.3758/s13414-014-0676-5.
Chambers, C.D., Mattingley, J.B., Moss, S.A., 2002. The octave illusion revisited: suppression or fusion between ears? J. Exp. Psychol. Hum. Percept. Perform. 28 (6), 1288–1302. https://doi.org/10.1037//0096-1523.28.6.1288.
Chambers, C.D., Mattingley, J.B., Moss, S.A., 2004. Reconsidering evidence for the suppression model of the octave illusion. Psychon. Bull. Rev. 11 (4), 642–666. https://doi.org/10.3758/BF03196617.
Chialvo, D.R., 2003. How we hear what is not there: a neural mechanism for the missing fundamental illusion. Chaos 13 (4), 1226–1230. https://doi.org/10.1063/1.1617771.
Chouinard, P.A., Royals, K.A., Landry, O., 2020. The strength of the Shepard illusion in children coincides with age and cognitive skills but not perceptual abilities. Attention, Percept., Psychophys. 82 (3), 1378–1390. https://doi.org/10.3758/s13414-019-01904-3.
Chowsilpa, S., Bamiou, D.-E., Koohi, N., 2021. Effectiveness of the auditory temporal ordering and resolution tests to detect central auditory processing disorder in adults with evidence of brain pathology: a systematic review and meta-analysis [Systematic Review]. Front. Neurol. 12. https://www.frontiersin.org/journals/neurology/articles/10.3389/fneur.2021.656117.
Conte, A., McGovern, E.M., Narasimham, S., Beck, R., Killian, O., O'Riordan, S., Reilly, R. B., Hutchinson, M., 2017. Temporal discrimination: mechanisms and relevance to adult-onset dystonia. Front. Neurol. 8, 625. https://doi.org/10.3389/fneur.2017.00625.
Deutsch, D., 1974. An auditory illusion. Nature 251 (5473), 307–309. https://doi.org/10.1038/251307a0.
Deutsch, D., 2019. Musical Illusions and Phantom Words: How Music and Speech Unlock Mysteries of the Brain. Oxford University Press.
Deutsch, D., Henthorn, T., Lapidis, R., 2011. Illusory transformation from speech to song. J. Acoust. Soc. Am. 129 (4), 2245–2252. https://doi.org/10.1121/1.3562174.
Deutsch, D., Roll, P.L., 1976. Separate "what" and "where" decision mechanisms in processing a dichotic tonal sequence. J. Exp. Psychol. Hum. Percept. Perform. 2 (1), 23–29. https://doi.org/10.1037//0096-1523.2.1.23.
Driver, J., 1996. Enhancement of selective listening by illusory mislocation of speech sounds due to lip-reading. Nature 381 (6577), 66–68. https://doi.org/10.1038/381066a0.
Falk, S., Rathcke, T., Dalla Bella, S., 2014. When speech sounds like music. J. Exp. Psychol.: Hum. Percept. Perform. 40, 1491–1506. https://doi.org/10.1037/a0036858.
Falk, S., Rathcke, T.V., 2010. On the speech-to-song illusion: evidence from German. In: Proceedings of the 5th Conference on Speech Prosody.
Frank, E.M., Jane, A.B., 2020. The Auditory System: Anatomy, Physiology, and Clinical Correlates, Second Edition. Plural Publishing, Inc (Vol. Second edition) [Book]. https://elib.tcd.ie/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=1885537.
Frankish, C., 2008. Precategorical acoustic storage and the perception of speech. J. Mem. Lang. 58 (3), 815–836. https://doi.org/10.1016/j.jml.2007.06.003.
Franosch, J.-M.P., Kempter, R., Fastl, H., van Hemmen, J.L., 2003. Zwicker Tone illusion and noise reduction in the auditory system. Phys. Rev. Lett. 90 (17), 178103. https://doi.org/10.1103/PhysRevLett.90.178103.
Franssen, N.V., 1962. Stereophony. English translation.
Frith, C.D., Frit, U., 1972. The solitaire illusion: an illusion of numerosity. Percept. Psychophys. 11 (6), 409–410. https://doi.org/10.3758/BF03206279.
Frost, R., Feldman, L.B., Katz, L., 1990. Phonological ambiguity and lexical ambiguity: effects on visual and auditory word recognition. J. Exp. Psychol.: Learn., Memory, Cogn. 16 (4), 569–580. https://doi.org/10.1037/0278-7393.16.4.569.
Frost, R., Repp, B.H., Katz, L., 1988. Can speech perception be influenced by simultaneous presentation of print? J. Mem. Lang. 27 (6), 741–755. https://doi.org/10.1016/0749-596X(88)90018-6.
Geiser, E., Gabrieli, J.D.E., 2013. Influence of rhythmic grouping on duration perception: a novel auditory illusion. PLoS One 8 (1), e54273. https://doi.org/10.1371/journal.pone.0054273.
Getz, L.M., Toscano, J.C., 2021. Rethinking the McGurk effect as a perceptual illusion. Attention, Percept., Psychophys. 83 (6), 2583–2598. https://doi.org/10.3758/s13414-021-02265-6.
Griffiths, T.D., Rees, A., Witton, C., Cross, P.M., Shakir, R.A., Green, G.G., 1997. Spatial and temporal auditory processing deficits following right hemisphere infarction. A

psychophysical study. Brain: J. Neurol. 120 (Pt 5), 785–794. https://doi.org/10.1093/brain/120.5.785.
Haas, V.H., 1951. Uber den Einfluss eines Einfachechos auf die Horsamkeit von Sprache. Acustica 1, 49–58.
Haas, V.H., 1972. The influence of a single echo on the audibility of speech. JAES 20 (2), 146–159.
Hari, R., 1995. Illusory directional hearing in humans. Neurosci. Lett. 189 (1), 29–30. https://doi.org/10.1016/0304-3940(95)11443-z.
Hartmann, W.M., Rakerd, B., 1989. Localization of sound in rooms IV: the Franssen effect. J. Acoust. Soc. Am. 86 (4), 1366–1373.
Hauser, M.D., McDermott, J., 2003. The evolution of the music faculty: a comparative perspective. Nat. Neurosci. 6 (7), 663–668. https://doi.org/10.1038/nn1080.
Hellström, Å. (2024). Chapter three - aesthetic valence: psychophysical perspectives. In T. D. Ben-Soussan, J. Glicksohn, & N. Srinivasan (Eds.), *Progress in Brain Research* (Vol. 287, pp. 45–70). Elsevier. 10.1016/bs.pbr.2024.05.005.
Hidaka, S., Teramoto, W., Sugita, Y., 2015. Spatiotemporal processing in crossmodal interactions for perception of the external world: a review. Front. Integr. Neurosci. 9. https://doi.org/10.3389/fnint.2015.00062.
Hirst, R.J., McGovern, D.P., Setti, A., Shams, L., Newell, F.N., 2020. What you see is what you hear: twenty years of research using the Sound-Induced Flash Illusion. Neurosci. Biobehav. Rev. 118, 759–774. https://doi.org/10.1016/j.neubiorev.2020.09.006.
Hubbard, T.L., 2014. The flash-lag effect and related mislocalizations: findings, properties, and theories. Psychol. Bull. 140 (1), 308–338. https://doi.org/10.1037/a0032899.
Hullfish, J., Sedley, W., Vanneste, S., 2019. Prediction and perception: insights for (and from) tinnitus. Neurosci. Biobehav. Rev. 102, 1–12. https://doi.org/10.1016/j.neubiorev.2019.04.008.
Jain, S., Nataraja, N.P., 2019. The relationship between temporal integration and temporal envelope perception in noise by males with mild sensorineural hearing loss. J. Int. Adv. Otol. 15 (2), 257–262. https://doi.org/10.5152/iao.2019.6555.
Jaisin, K., Suphanchaimat, R., Figueroa Candia, M.A., Warren, J.D., 2016. The speech-to-song illusion is reduced in speakers of tonal (vs. Non-Tonal) languages. Front. Psychol. 7, 662. https://doi.org/10.3389/fpsyg.2016.00662.
Kafadar, E., Fisher, V.L., Quagan, B., Hammer, A., Jaeger, H., Mourgues, C., Thomas, R., Chen, L., Imtiaz, A., Sibarium, E., Negreira, A.M., Sarisik, E., Polisetty, V., Benrimoh, D., Sheldon, A.D., Lim, C., Mathys, C., Powers, A.R., 2022. Conditioned hallucinations and prior overweighting are state-sensitive markers of hallucination susceptibility. Biol. Psychiatry 92 (10), 772–780. https://doi.org/10.1016/j.biopsych.2022.05.007.
Kaliuzhna, M., Stein, T., Sterzer, P., Seymour, K.J., 2020. Examining motion speed processing in schizophrenia using the flash lag illusion. Schizophrenia Res.: Cognition 19, 100165. https://doi.org/10.1016/j.scog.2019.100165.
Kamke, M.R., Vieth, H.E., Cottrell, D., Mattingley, J.B., 2012. Parietal disruption alters audiovisual binding in the sound-induced flash illusion. Neuroimage 62 (3), 1334–1341. https://doi.org/10.1016/j.neuroimage.2012.05.063.
Kammers, M.P.M., de Vignemont, F., Verhagen, L., Dijkerman, H.C., 2009. The rubber hand illusion in action. Neuropsychologia 47 (1), 204–211. https://doi.org/10.1016/j.neuropsychologia.2008.07.028.
Kidd, J.C., Hogben, J.H., 2004. Quantifying the auditory saltation illusion: an objective psychophysical methodology. J. Acoust. Soc. Am. 116 (2), 1116–1125. https://doi.org/10.1121/1.1766054.
Kitajima, N., Yamashita, Y., 1999. Dynamic capture of sound motion by light stimuli moving in three-dimensional space. Percept. Mot. Skills. 89 (3 Pt 2), 1139–1158. https://doi.org/10.2466/pms.1999.89.3f.1139.
Komatsu, H., 2006. The neural mechanisms of perceptual filling-in. Nat. Rev. Neurosci. 7 (3), 220–231. https://doi.org/10.1038/nrn1869.
Kwakye, L.D., Foss-Feig, J.H., Cascio, C.J., Stone, W.L., Wallace, M.T., 2011. Altered auditory and multisensory temporal processing in autism spectrum disorders [Original Research]. Front. Integr. Neurosci. 4. https://www.frontiersin.org/journals/integrative-neuroscience/articles/10.3389/fnint.2010.00129.
Kwon, S.-H., Park, J.K., Koh, Y.H., 2023. A systematic review and meta-analysis on the effect of virtual reality-based rehabilitation for people with Parkinson's disease. J. Neuroeng. Rehabil. 20 (1), 94. https://doi.org/10.1186/s12984-023-01219-3.
Leske, S., Tse, A., Oosterhof, N.N., Hartmann, T., Müller, N., Keil, J., Weisz, N., 2014. The strength of alpha and beta oscillations parametrically scale with the strength of an illusory auditory percept. Neuroimage 88, 69–78. https://doi.org/10.1016/j.neuroimage.2013.11.014.
Letowski, T.R., Letowski, S.T., 2012. Auditory spatial perception: auditory localization. Army Research Laboratory Aberdeen Proving Ground MD Human Research and Engineering Directorate.
Licklider, J.C.R., 1954. Periodicity" pitch and "place" pitch. J. Acoust. Soc. Am. 26 (5_Supplement), 945. https://doi.org/10.1121/1.1928005.
Machado Sotomayor, M.J., Arufe-Giráldez, V., Ruíz-Rico, G., Navarro-Patón, R., 2021. Music therapy and Parkinson's disease: a systematic review from 2015 to 2020. Int. J. Environ. Res. Public Health 18 (21). https://doi.org/10.3390/ijerph182111618.
Margulis, E.H., Simchy-Gross, R., 2016. Repetition enhances the musicality of randomly generated tone sequences. Music Percept. 33 (4), 509–514. https://doi.org/10.1525/mp.2016.33.4.509.
Margulis, E.H., Simchy-Gross, R., Black, J.L., 2015. Pronunciation difficulty, temporal regularity, and the speech-to-song illusion. Front. Psychol. 6. https://doi.org/10.3389/fpsyg.2015.00048.
Marquez-Garcia, A.V., Magnuson, J., Morris, J., Iarocci, G., Doesburg, S., Moreno, S., 2022. Music therapy in autism spectrum disorder: a systematic review. Rev. J. Autism. Dev. Disord. 9 (1), 91–107. https://doi.org/10.1007/s40489-021-00246-x.
Massaro, D.W., Cohen, M.M., 1983. Categorical or continuous speech perception: a new test. Speech. Commun. 2 (1), 15–35. https://doi.org/10.1016/0167-6393(83)90061-4.
Mateeff, S., Hohnsbein, J., Noack, T., 1985. Dynamic visual capture: apparent auditory motion induced by a moving visual target. Perception 14 (6), 721–727. https://doi.org/10.1068/p140721.
Mc Govern, E.M., O'Connor, E., Beiser, I., Williams, L., Butler, J.S., Quinlivan, B., Narasimham, S., Beck, R., Reilly, R.B., O'Riordan, S., Hutchinson, M., 2017. Menstrual cycle and the temporal discrimination threshold. Physiol. Meas. 38 (2), N65–n72. https://doi.org/10.1088/1361-6579/38/2/n65.
Mcbeath, M.K., Neuhoff, J.G., 2002. The Doppler effect is not what you think it is: dramatic pitch change due to dynamic intensity change. Psychon. Bull. Rev. 9 (2), 306–313. https://doi.org/10.3758/BF03196286.
McGurk, H., MacDonald, J., 1976. Hearing lips and seeing voices. Nature 264 (5588), 746–748. https://doi.org/10.1038/264746a0.
McWalter, R., McDermott, J.H., 2019. Illusory sound texture reveals multi-second statistical completion in auditory scene analysis. Nat. Commun. 10 (1), 5096. https://doi.org/10.1038/s41467-019-12893-0.
Miller, G.A., Licklider, J.C.R., 1950. The Intelligibility of Interrupted Speech. J. Acoust. Soc. Am. 22 (2), 167–173. https://doi.org/10.1121/1.1906584.
Mohan, A., Bhamoo, N., Riquelme, J.S., Long, S., Norena, A., Vanneste, S., 2020. Investigating functional changes in the brain to intermittently induced auditory illusions and its relevance to chronic tinnitus. Hum. Brain Mapp. 41 (7), 1819–1832. https://doi.org/10.1002/hbm.24914.
Møller, C., Højlund, A., Bærentsen, K.B., Hansen, N.C., Skewes, J.C., Vuust, P., 2018. Visually induced gains in pitch discrimination: linking audio-visual processing with auditory abilities. Attention, Percept., Psychophys. 80 (4), 999–1010. https://doi.org/10.3758/s13414-017-1481-8.
Møller, H., Sørensen, M.F., Hammershøi, D., Jensen, C.B., 1995. Head-related transfer functions of human subjects. J. Audio Eng. Soc. 43 (5), 300–321.
Müller, N., Keil, J., Obleser, J., Schulz, H., Grunwald, T., Bernays, R.-L., Huppertz, H.-J., Weisz, N., 2013. You can't stop the music: reduced auditory alpha power and coupling between auditory and memory regions facilitate the illusory perception of music during noise. Neuroimage 79, 383–393. https://doi.org/10.1016/j.neuroimage.2013.05.001.
Mullin, H.A.C., Norkey, E.A., Kodwani, A., Vitevitch, M.S., Castro, N., 2021. Does age affect perception of the speech-to-song illusion? PLoS One 16 (4), e0250042. https://doi.org/10.1371/journal.pone.0250042.
Nakajima, Y., Sasaki, T., Kanafuka, K., Miyamoto, A., Remijn, G., Ten Hoopen, G., 2000. Illusory recouplings of onsets and terminations of glide tone components. Percept. Psychophys. 62 (7), 1413–1425. https://doi.org/10.3758/BF03212143.
Nakajima, Y., Ten Hoopen, G., Sasaki, T., Yamamoto, K., Kadota, M., Simons, M., Suetomi, D., 2004. Time-shrinking: the process of unilateral temporal assimilation. Perception. 33 (9), 1061–1079.
Neuhoff, J.G., McBeath, M.K., 1996. The Doppler illusion: the influence of dynamic intensity change on perceived pitch. J. Exp. Psychol.: Hum. Percept. Perform. 22 (4), 970–985. https://doi.org/10.1037/0096-1523.22.4.970.
Nijhawan, R., 1994. Motion extrapolation in catching. Nature 370 (6487), 256–257. https://doi.org/10.1038/370256b0.
Norena, A., Micheyl, C., Chery-Croze, S., 2000. An auditory negative after-image as a human model of tinnitus. Hear. Res. 149 (1–2), 24–32. https://doi.org/10.1016/s0378-5955(00)00158-1.
Noreña, A.J., Eggermont, J.J., 2003. Neural correlates of an auditory afterimage in primary auditory cortex. J. Assoc. Res. Otolaryngol. 4 (3), 312–328. https://doi.org/10.1007/s10162-002-3039-1.
O'Donoghue, J., Moss, H., Clements-Cortes, A., Freeley, C., 2020. Therapist and individual experiences and perceptions of music therapy for adolescents who stutter: a qualitative exploration. Nord. J. Music. Ther. 29 (4), 353–370. https://doi.org/10.1080/08098131.2020.1745872.
Oxenham, A.J., 2008. Pitch perception and auditory stream segregation: implications for hearing loss and cochlear implants. Trends. Amplif. 12 (4), 316–331. https://doi.org/10.1177/1084713808325881.
Oxenham, A.J., 2012. Pitch perception. J. Neurosci. 32 (39), 13335–13338. https://doi.org/10.1523/jneurosci.3815-12.2012.
Oxenham, A.J., 2018. How we hear: the perception and neural coding of sound. Annu. Rev. Psychol. 69, 27–50. https://doi.org/10.1146/annurev-psych-122216-011635.
Pantev, C., Hoke, M., Lütkenhöner, B., Lehnertz, K., 1989. Tonotopic organization of the auditory cortex: pitch versus frequency representation. Science 246 (4929), 486–488. https://doi.org/10.1126/science.2814476.
Pantev, C., Wollbrink, A., Roberts, L.E., Engelien, A., Lütkenhöner, B., 1999. Short-term plasticity of the human auditory cortex. Brain Res. 842 (1), 192–199. https://doi.org/10.1016/s0006-8993(99)01835-1.
Parise, C.V., Knorre, K., Ernst, M.O., 2014. Natural auditory scene statistics shapes human spatial hearing. Proc. Natl. Acad. Sci. 111 (16), 6104–6108. https://doi.org/10.1073/pnas.1322705111.
Pearl, D., Yodashkin-Porat, D., Katz, N., Valevski, A., Aizenberg, D., Sigler, M., Weizman, A., Kikinzon, L., 2009. Differences in audiovisual integration, as measured by McGurk phenomenon, among adult and adolescent patients with schizophrenia and age-matched healthy control groups. Compr. Psychiatry 50 (2), 186–192. https://doi.org/10.1016/j.comppsych.2008.06.004.
Phillips, D.P., Hall, S.E., 2001. Spatial and temporal factors in auditory saltation. J. Acoust. Soc. Am. 110 (3 Pt 1), 1539–1547. https://doi.org/10.1121/1.1396329.
Pinker, S., 2005. So how does the mind work? Mind. Lang. 20 (1), 1–24.
Plack, C., Fay, R., Oxenham, A., & Popper, A. (2005). *Pitch: neural coding and perception.*

Powers, A.R., Mathys, C., Corlett, P.R., 2017. Pavlovian conditioning–induced hallucinations result from overweighting of perceptual priors. Science 357 (6351), 596–600. https://doi.org/10.1126/science.aan3458.
Prpic, V., Luccio, R., 2016. Do you hear more piano or drum sounds? an auditory version of the solitaire illusion. Perception (12), 45. https://doi.org/10.1177/0301006616672580.
Radeau, M., Bertelson, P., 1974. The after-effects of ventriloquism. Q. J. Exp. Psychol. 26 (1), 63–71. https://doi.org/10.1080/14640747408400388.
Rammsayer, T.H., Buttkus, F., Altenmüller, E., 2012. Musicians do better than nonmusicians in both auditory and visual timing tasks. Music Percept.: Interdiscip. J. 30 (1), 85–96.
Ravignani, A., Madison, G., 2017. The paradox of isochrony in the evolution of human rhythm. Front. Psychol. 8, 1820. https://doi.org/10.3389/fpsyg.2017.01820.
Remijn, G.B., Nakajima, Y., 2005. The perceptual integration of auditory stimulus edges: an illusory short tone in stimulus patterns consisting of two partly overlapping glides. J. Exp. Psychol. Hum. Percept. Perform. 31 (1), 183–192. https://doi.org/10.1037/0096-1523.31.1.183.
Repp, B.H., Frost, R., Zsiga, E., 1992. Lexical mediation between sight and sound in speechreading. Q. J. Exp. Psychol. Sect. A 45 (1), 1–20. https://doi.org/10.1080/14640749208401313.
Rhode, W.S., Greenberg, S., 1994. Lateral suppression and inhibition in the cochlear nucleus of the cat. J. Neurophysiol. 71 (2), 493–514. https://doi.org/10.1152/jn.1994.71.2.493.
Riecke, L., Van Orstal, A.J., Formisano, E., 2008. The auditory continuity illusion: a parametric investigation and filter model. Percept. Psychophys. 70 (1), 1–12. https://doi.org/10.3758/pp.70.1.1.
Rieger, K.L., Chernomas, W.M., McMillan, D.E., Morin, F.L., Demczuk, L., 2016. Effectiveness and experience of arts-based pedagogy among undergraduate nursing students: a mixed methods systematic review. JBI Evid. Synth. 14 (11), 139–239.
Rimmele, J.M., Sussman, E., Poeppel, D., 2015. The role of temporal structure in the investigation of sensory memory, auditory scene analysis, and speech perception: a healthy-aging perspective. Int. J. Psychophysiol. 95 (2), 175–183. https://doi.org/10.1016/j.ijpsycho.2014.06.010.
Roberts Richard, A., Lister Jennifer, J., 2004. Effects of age and hearing loss on gap detection and the precedence effect. J. Speech, Lang., Hear. Res. 47 (5), 965–978. https://doi.org/10.1044/1092-4388(2004/071.
Sasaki, T., Nakajima, Y., Hoopen, G.T., Van Buuringen, E., Massier, B., Kojo, T., Kuroda, T., Ueda, K., 2010. Time stretching: illusory lengthening of filled auditory durations. Attention, Percept., Psychophys. 72 (5), 1404–1421. https://doi.org/10.3758/APP.72.5.1404.
Scharine, A.A., Letowski, T.R., 2009. Auditory conflicts and illusions. Helmet-Mounted Displays: Sensation, Perception and Cognition Issues, pp. 579–598.
Schilling, A., Tziridis, K., Schulze, H., & Krauss, P. (2021). Chapter 6 - the stochastic resonance model of auditory perception: a unified explanation of tinnitus development, Zwicker Tone illusion, and residual inhibition. In B. Langguth, T. Kleinjung, D. De Ridder, W. Schlee, & S. Vanneste (Eds.), *Progress in Brain Research* (Vol. 262, pp. 139–157). Elsevier.
Schlaug, G., Norton, A., Marchina, S., Zipse, L., Wan, C.Y., 2010. From singing to speaking: facilitating recovery from nonfluent aphasia. Future Neurol. 5 (5), 657–665. https://doi.org/10.2217/fnl.10.44.
Senna, I., Parise, C.V., Ernst, M.O., 2015. Hearing in slow-motion: humans underestimate the speed of moving sounds. Sci. Rep. 5 (1), 14054. https://doi.org/10.1038/srep14054.
Senna, I., Parise, C.V., Ernst, M.O., 2017. Modulation frequency as a cue for auditory speed perception. Proc. Biol. Sci. 284 (1858), 20170673. https://doi.org/10.1098/rspb.2017.0673.
Shamma, S., 2001. On the role of space and time in auditory processing. Trends Cogn. Sci. 5 (8), 340–348. https://doi.org/10.1016/S1364-6613(00)01704-6.
Shepard, R.N., 1964. Circularity in judgments of relative pitch. J. Acoust. Soc. Am. 36 (12), 2346–2353. https://doi.org/10.1121/1.1919362.
Shinn, J.B., 2003. Temporal processing: the basics. Hear. J. (7), 56. In: https://journals.lww.com/thehearingjournal/fulltext/2003/07000/temporal_processing__the_basics.10.aspx.
Shore, D.I., Hall, S.E., Klein, R.M., 1998. Auditory saltation: a new measure for an old illusion. J. Acoust. Soc. Am. 103 (6), 3730–3733. https://doi.org/10.1121/1.423093.
Simchy-Gross, R., Margulis, E.H., 2018. The sound-to-music illusion: repetition can musicalize nonspeech sounds. Music. Sci. 1. https://doi.org/10.1177/2059204317731992, 2059204317731992.
Soto-Faraco, S., Lyons, J., Gazzaniga, M., Spence, C., Kingstone, A., 2002. The ventriloquist in motion: illusory capture of dynamic information across sensory modalities. Brain Res. Cogn. Brain Res. 14 (1), 139–146. https://doi.org/10.1016/s0926-6410(02)00068-x.
Spelke, E.S., Phillips, A., Woodward, A.L., 1995. Infants' knowledge of object motion and human action. Causal Cognition: A Multidisciplinary Debate. Clarendon Press/Oxford University Press, pp. 44–78.
Spillmann, L., Otte, T., Hamburger, K., Magnussen, S., 2006. Perceptual filling-in from the edge of the blind spot. Vision Res. 46 (25), 4252–4257. https://doi.org/10.1016/j.visres.2006.08.033.
Stein, B.E., Meredith, M.A., 1993. The Merging of the Senses. The MIT Press.
Stein, B.E., Stanford, T.R., 2008. Multisensory integration: current issues from the perspective of the single neuron. Nat. Rev. Neurosci. 9 (4), 255–266. https://doi.org/10.1038/nrn2331.
Su, Y.-H., 2014. Visual enhancement of auditory beat perception across auditory interference levels. Brain Cogn. 90, 19–31. https://doi.org/10.1016/j.bandc.2014.05.003.
Thurlow, W.R., Jack, C.E., 1973. Certain Determinants of the "Ventriloquism Effect. Percept. Mot. Skills. 36 (3_suppl), 1171–1184. https://doi.org/10.2466/pms.1973.36.3c.1171.
Tiehen, J., 2023. Perception as controlled hallucination. Anal. Philos. 64 (4), 355–372. https://doi.org/10.1111/phib.12268.
Tierney, A., Dick, F., Deutsch, D., Sereno, M., 2013. Speech versus song: multiple pitch-sensitive areas revealed by a naturally occurring musical illusion. Cereb. Cortex 23 (2), 249–254. https://doi.org/10.1093/cercor/bhs003.
Väljamäe, A., 2009. Auditorily-induced illusory self-motion: a review. Brain Res. Rev. 61 (2), 240–255. https://doi.org/10.1016/j.brainresrev.2009.07.001.
Vesely, S., Dolci, C., & Dolci, S. (2018). *The Doppler effect in the perception of sound* (Vol. 1990).
Vroomen, J., de Gelder, B., 2003. Visual motion influences the contingent auditory motion aftereffect. Psychol. Sci. 14 (4), 357–361. https://doi.org/10.1111/1467-9280.24431.
Wallace, M.T., Woynaroski, T.G., Stevenson, R.A., 2020. Multisensory integration as a window into orderly and disrupted cognition and communication. Annu. Rev. Psychol. 71, 193–219. https://doi.org/10.1146/annurev-psych-010419-051112.
Wallach, H., Newman, E.B., Rosenzweig, M.R., 1949. The precedence effect in sound localization. Am. J. Psychol. 62 (3), 315–336. https://doi.org/10.2307/1418275.
Ward, K.S., 2013. Creative arts-based pedagogies in early childhood Education for Sustainability (EfS): challenges and possibilities. Australian J. Environ. Educ. 29 (2), 165–181. https://doi.org/10.1017/aee.2014.4.
Warren, R.M., 1970. Perceptual restoration of missing speech sounds. Science 167 (3917), 392–393. https://doi.org/10.1126/science.167.3917.392.
Warren, R.M., Obusek, C.J., Ackroff, J.M., 1972. Auditory induction: perceptual synthesis of absent sounds. Science 176 (4039), 1149–1151. https://doi.org/10.1126/science.176.4039.1149.
Warren, R.M., Wrightson, J.M., Puretz, J., 1988. Illusory continuity of tonal and infratonal periodic sounds. J. Acoust. Soc. Am. 84 (4), 1338–1342. https://doi.org/10.1121/1.396632.
Wertheimer, M., 1923. Untersuchungen zur Lehre von der Gestalt. II. Psychol. Forsch. 4 (1), 301–350. https://doi.org/10.1007/BF00410640.
Wertheimer, M. (1938). Laws of organization in perceptual forms.
White, T.P., Wigton, R.L., Joyce, D.W., Bobin, T., Ferragamo, C., Wasim, N., Lisk, S., Shergill, S.S., 2014. Eluding the illusion? Schizophrenia, dopamine and the McGurk effect. Front. Hum. Neurosci. 8, 565. https://doi.org/10.3389/fnhum.2014.00565.
Williams, L.J., Butler, J.S., Molloy, A., McGovern, E., Beiser, I., Kimmich, O., Quinlivan, B., O'Riordan, S., Hutchinson, M., Reilly, R.B., 2015. Young women do it better: sexual dimorphism in temporal discrimination. Front. Neurol. 6, 160. https://doi.org/10.3389/fneur.2015.00160.
Yalachkov, Y., Bergmann, H.J., Soydaş, D., Buschenlange, C., Fadai Motlagh, L.Y., Naumer, M.J., Kaiser, J., Frisch, S., Behrens, M., Foerch, C., Gehrig, J., 2019. Cognitive impairment in multiple sclerosis is reflected by increased susceptibility to the sound-induced flash illusion [Original Research]. Front. Neurol. 10. https://doi.org/10.3389/fneur.2019.00373.
Yasoda-Mohan, A., Chen, F., Ó Sé, C., Allard, R., Ost, J., Vanneste, S., 2024. Phantom perception as a Bayesian inference problem: a pilot study. J. Neurophysiol. 131 (6), 1311–1327. https://doi.org/10.1152/jn.00349.2023.
Yordamlı, A., Erdener, D., 2018. Auditory–visual speech integration in bipolar disorder: a preliminary study. Languages 3 (4).
Yoshida, S., Hase, K., Heim, O., Kobayasi, K.I., Hiryu, S., 2024. Doppler detection triggers instantaneous escape behavior in scanning bats. iScience 27 (3), 109222. https://doi.org/10.1016/j.isci.2024.109222.
Zaidel, D.W., 2020. The arts and human symbolic cognition: art is for social communication. The Cambridge Handbook of the Imagination. Cambridge University Press, pp. 593–607. https://doi.org/10.1017/9781108580298.036.
Zeljic, K., Xiang, Q., Wang, Z., Pan, Y., Shi, Y., Zhou, Z., Wang, Z., Liu, D., 2021. Heightened perception of illusory motion is associated with symptom severity in schizophrenia patients. Prog. Neuro-Psychopharmacol. Biol. Psychiatry 104, 110055. https://doi.org/10.1016/j.pnpbp.2020.110055.
Zwicker, E., 1964. Negative afterimage" in hearing. J. Acoust. Soc. Am. 36 (12), 2413–2415. https://doi.org/10.1121/1.1919373.