CORTEX 103 (2018) 130–141

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research Report

Sounds from seeing silent motion: Who hears them, and what looks loudest?

Christopher J. Fassnidge and Elliot D. Freeman[*]

Cognitive Neuroscience Research Unit, City University of London, UK

ARTICLE INFO

Article history:
Received 24 October 2017
Reviewed 5 December 2017
Revised 16 January 2018
Accepted 23 February 2018
Action editor Jason Mattingley
Published online 9 March 2018

Keywords:
Individual differences
Audiovisual perception
Synaesthesia

ABSTRACT

Some people hear what they see: car indicator lights, flashing neon shop signs, and people's movements as they walk may all trigger an auditory sensation, which we call the visual-evoked auditory response (vEAR or 'visual ear'). We have conducted the first large-scale online survey (N > 4000) of this little-known phenomenon. We analysed the prevalence of vEAR, what induces it, and what other traits are associated with it.

We assessed prevalence by asking whether respondents had previously experienced vEAR. Participants then rated silent videos for vividness of evoked auditory sensations, and answered additional trait questions.

Prevalence appeared higher relative to other typical synaesthesias. Prior awareness and video ratings were associated with greater frequency of other synaesthesias, including flashes evoked by sounds, and musical imagery. Higher-rated videos often depicted meaningful events that predicted sounds (e.g., collisions). However, even videos containing abstract flickering or moving patterns could also elicit higher ratings, despite having no predictable association with sounds. Such videos had higher levels of raw 'motion energy' (ME), which we quantified using a simple computational model of motion processing in early visual cortex. Critically, only respondents reporting prior awareness of vEAR tended to show a positive correlation between video ratings and ME.

This specific sensitivity to ME suggests that in vEAR, signals from visual motion processing may affect audition relatively directly without requiring higher-level interpretative processes. Our other findings challenge the popular assumption that individuals with synaesthesia are rare and have ideosyncratic patterns of brain hyper-connectivity. Instead, our findings of apparently high prevalence and broad associations with other synaesthesias and traits are jointly consistent with a common dependence on normal variations in physiological mechanisms of disinhibition or excitability of sensory brain areas and their functional connectivity. The prevalence of vEAR makes it easier to test such hypotheses further, and makes the results more relevant to understanding not only synaesthetic anomalies but also normal perception.

---

[*] *Corresponding author*. City, University of London, Northampton Square, London, EC1V 0HB, UK.
E-mail address: elliot.freeman@city.ac.uk (E.D. Freeman).

https://doi.org/10.1016/j.cortex.2018.02.019

1. Introduction

It is well known that what we see can influence what we hear. For example the sight of a person's lip movements can enhance speech comprehension or even change our interpretation of the speech sounds (McGurk & MacDonald, 1976; Sumby & Pollack, 1954), while the movements of a musician can dominate the perceived quality of their performance even among expert listeners (Tsay, 2013). However it is much less appreciated that vision not only modulates perception of concurrent auditory stimuli, but sometimes can also induce the perception of new auditory sensations (Saenz & Koch, 2008). For example, some individuals claim they can 'hear' flashing car indicator lights or shop displays, or people's movements while walking or speaking. Over the last decade since Saenz and Koch's (2008) short report, there has been remarkably little further research on this intriguing phenomenon (Fassnidge, Cecconi Marcotti, & Freeman, 2017; Rothen, Bartl, Franklin, & Ward, 2017), which we call the 'visually-evoked auditory response' (vEAR, also known as 'hearing motion synaesthesia'). Many questions therefore remain to be answered. For example, it is currently unknown whether vEAR represents a form of high-level cognitive association or imagery, or a genuine form of synaesthesia-like sensory crosstalk. It is also unknown how prevalent vEAR is in a large population, what traits characterise people who experience it, and whether vEAR depends on normal or abnormal brain mechanisms. We address these questions here using a large-scale internet survey. We have assessed the prevalence of vEAR, and for the first time analysed both the kinds of visual motion stimuli that evoke high ratings of auditory sensations, and the individual traits that predict high ratings and vEAR susceptibility.

Our study is relevant to two ongoing debates about the neural underpinnings of synaesthesia (Hubbard & Ramachandran, 2005). One debate concerns the neuro-architecture that results in sensory cross-talk: whether synaesthesia is mediated by feedback from high-level semantic representations, or whether it involves relatively more direct cross-wiring between sensory modalities. For example, some have argued in favour of direct anatomical connections between brain areas (Ramachandran & Hubbard, 2001), while other evidence suggests that this phenomenon requires prior semantic interpretation of the inducing stimulus (Mattingley, Rich, Yelland, & Bradshaw, 2001; Myles, Dixon, Smilek, & Merikle, 2003; Smilek, Dixon, Cudahy, & Merikle, 2001). A second debate concerns the neurophysiological causes of synaesthesia: it might result from unusual anatomical patterns of connectivity (Bargary & Mitchell, 2008; Baron-Cohen, 1996; Hubbard & Ramachandran, 2005; Tomson et al., 2011), or from more systemic physiological variables which disinhibit normally-occurring connections between sensory areas or render the areas themselves more excitable (Grossenbacher & Lovelace, 2001; Neufeld et al., 2012).

Concerning neuro-architecture, the role of semantic representations can be ambiguous in some synaesthesias which involve inducers that are relatively high-level or cultural in origin, such as letters or words evoking colours (Bor, Rothen, Schwartzman, Clayton, & Seth, 2014; Witthoft, Winawer, & Eagleman, 2015). There is potentially less ambiguity where synaesthesia involves associations between more basic sensory dimensions such as sound and colour, or in this case between visual movement and sound. However in vEAR there may still be two routes to inducing a sound: one which depends on relatively direct crosstalk between areas processing low-level visual motion cues and audition, and a higher-level route that depends on prior semantic analysis of the visual scene and predictions about whether the depicted events are likely to produce sounds, associated for example with friction, collisions or explosions. Fortunately, in vEAR it is straightforward to quantify how much low-level 'motion energy' (ME) is present in the image sequence as the patterns of luminance change over space and time, for example using a computational approach which models the spatiotemporal sensitivity of cells in early visual cortex to moving patterns (Adelson & Bergen, 1985). The contribution of this ME to the perception of vEAR may be measured independently from the higher-level semantic content of the images.

Concerning the second debate about the neurophysiological causes of synaesthesia, the assumption that synaesthesias depend on abnormal neural connectivity is supported by evidence that typical synaesthesias are both fairly rare and highly idiosyncratic (Simner et al., 2006; Ward, 2013). However another reason for the apparent rarity of synaesthesia might be that it is unusual to regularly encounter in nature the specific combinations of inducers and concurrents that are typically associated in synaesthesia (Fassnidge et al., 2017). For example, consistent pairings between specific letters and colours do not occur in the natural environment, although grapheme-colour associations may be reinforced following repeated exposure in childhood to coloured fridge magnets (Bor et al., 2014; Witthoft et al., 2015). On this statistical view, one might expect higher prevalence of vEAR than other synaesthesias because vision and audition are naturally highly correlated with each other, for example every time we see and hear a person talking or making a footstep. In support of this, we previously found that 22% of our lab noticed faint sounds evoked by silent 'Morse-code' flashes (Fassnidge et al., 2017), however this was a small sample ($N = 37$). If vEAR were found to have greater prevalence than other typical synaesthesias, this would suggest that some synaesthesia-like phenomena could occur via relatively normal rather than rarely occurring patterns of neural connectivity.

Despite the individual idiosyncrasy of synaesthetic associations, there is evidence that people with one kind of synaesthesia are more likely to report others (Barnett et al., 2008; Rothen et al., 2013; Sagiv, Simner, Collins, Butterworth, & Ward, 2006), plus evidence of distinct personality profiles (Banissy et al., 2013; Rouw & Scholte, 2016), comorbidities for example with schizotypy and autism spectrum disorders (Banissy et al., 2012; Baron-Cohen et al., 2013; Ward et al., 2017), and other generalised benefits in sensory acuity (Banissy, Walsh, & Ward, 2009). Such broad patterns of association would be more supportive of the notion that there are systemic variables, possibly of a genetic origin, governing the expression of synaesthetic phenomena and its associated traits (Barnett et al., 2008; Carmichael & Simner, 2013). Such variables might broadly affect development of connectivity, and/or impact on cortical excitability, which has independently been linked to