Cognition 134 (2015) 245–251

Contents lists available at ScienceDirect

Cognition

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

Brief article

Evidence for a supra-modal representation of emotion from cross-modal adaptation

Annie Pye, Patricia E.G. Bestelmeyer *

School of Psychology, Bangor University, Gwynedd LL57 2AS, UK

ARTICLE INFO

Article history:
Received 31 July 2013
Revised 27 August 2014
Accepted 2 November 2014
Available online 19 November 2014

Keywords:
Supra-modal representation
Cross-modal
Adaptation
Emotion
Voice

ABSTRACT

Successful social interaction hinges on accurate perception of emotional signals. These signals are typically conveyed multi-modally by the face and voice. Previous research has demonstrated uni-modal contrastive aftereffects for emotionally expressive faces *or* voices. Here we were interested in whether these aftereffects transfer across modality as theoretical models predict. We show that adaptation to facial expressions elicits significant auditory aftereffects. Adaptation to angry facial expressions caused ambiguous vocal stimuli drawn from an anger-fear morphed continuum to be perceived as less angry and more fearful relative to adaptation to fearful faces. In a second experiment, we demonstrate that these aftereffects are not dependent on learned face-voice congruence, i.e. adaptation to one facial identity transferred to an unmatched voice identity. Taken together, our findings provide support for a supra-modal representation of emotion and suggest further that identity and emotion may be processed independently from one another, at least at the supra-modal level of the processing hierarchy.

1. Introduction

The face and voice convey important information about an individual such as gender, identity or emotional state. The accurate perception and integration of relevant signals from the face and voice plays an important role in successful social interaction. The 'auditory face' model of voice perception (Belin, Fecteau, & Bedard, 2004), based on the model of face perception by Bruce and Young (1986), suggests that following separate low-level visual and auditory analysis, (lip) speech, affect and identity are processed by dissociable perceptual systems. Furthermore, the model hypothesises the presence of interactions between face and voice at each of the dissociable processing pathways. While Belin et al.'s model was published nearly a decade ago and is widely cited in the literature, only few studies have empirically tested its assumptions. One method of exploring the predictions made by this model is with the use of adaptation paradigms.

Adaptation refers to the process during which continued stimulation results in biased perception towards opposite features of the adapting stimulus (Grill-Spector et al., 1999). The use of adaptation paradigms provides information about what is encompassed within a given representation (Dosenbach et al., 2007) and enables further understanding of the neural underpinnings of our sensory and perceptual experiences through the study of aftereffects (Kohn, 2007). Research using adaptation has revealed neural populations tuned to respond to specific stimulus attributes by isolating and subsequently distorting the perception of these attributes (Grill-Spector et al., 1999; Winston, Henson, Fine-Goulden, & Dolan, 2004).

* Corresponding author at: School of Psychology, Bangor University, Brigantia Building, Penrallt Road, Bangor LL57 2AS, UK. Tel.: +44 (0) 1248 383488; fax: +44 (0) 1248 382599.

E-mail address: p.bestelmeyer@bangor.ac.uk (P.E.G. Bestelmeyer).

http://dx.doi.org/10.1016/j.cognition.2014.11.001
0010-0277/

Adaptation is therefore a useful tool to investigate different levels of processing including "low-level" stimulus properties such as motion, shape or colour (e.g. Wright, 1934) but also "higher-level" properties such as facial features. Webster and MacLin (1999) were the first to show that adaptation to a face with distorted features resulted in a perceptual bias in which the test face was interpreted as being distorted in the opposite direction to that of the adaptor face. Aftereffects have been found for a variety of facial features and have been widely used to investigate the mechanisms that underpin face perception (e.g. Bestelmeyer, Jones, DeBruine, Little, & Welling, 2010; Bestelmeyer et al., 2008; Fox & Barton, 2007; Jenkins, Beaver, & Calder, 2006; Leopold, O'Toole, Vetter, & Blanz, 2001; Leopold, Rhodes, Muller, & Jeffery, 2005; Little, DeBruine, & Jones, 2005; Webster, Kaping, Mizokami, & Duhamel, 2004; Welling et al., 2009). While this research has largely concentrated on evaluating face perception models it has recently been shown that corresponding auditory aftereffects can be elicited for voices (e.g. identity (Zäske, Schweinberger, & Kawahara, 2010), gender (Schweinberger et al., 2008), age (Zäske & Schweinberger, 2011) and affect (Bestelmeyer, Rouger, DeBruine, & Belin, 2010; Skuk & Schweinberger, 2013)).

Few cross-modal adaptation studies have probed the existence of supra-modal, or modality-independent, representations of person-specific information such as identity. While Belin et al. (2004) clearly predict cross-modal interactions between perceptual pathways (i.e. lip speech, identity, affect) empirical studies have yielded inconsistent results. For example, studies concerning cross-modal gender adaptation between face and voice have usually failed to find significant aftereffects (Owlia & Jordan, 2009; Schweinberger et al., 2008; but see Little, Feinberg, DeBruine, & Jones, 2013). While cross-modal studies regarding identity information unanimously report aftereffects (Hills, Elward, & Lewis, 2010; Zäske et al., 2010), cross-modal studies of emotion again report mixed findings. Adapting to sentences spoken in an emotional tone did not distort perception of static facial expressions (Fox & Barton, 2007), yet when using dynamic facial adaptors cross-modal aftereffects were found for emotional prosody (expressed as brief syllables) but only in male listeners (Skuk & Schweinberger, 2013). Experiment 1 explored whether adaptation to facial expressions affects the perception of non-linguistic, vocal expressions.

An on-going debate exists as to whether functionally different aspects of faces are indeed processed independently of one another, as several theoretical models suggest (Belin, Bestelmeyer, Latinus, & Watson, 2011; Belin et al., 2004; Bruce & Young, 1986; Haxby, Hoffman, & Gobbini, 2000; Le Gal & Bruce, 2002; Young, 1998; but see Young & Bruce, 2011 for a revised account). This idea implies that the processing of one dimension (e.g. expression) is not influenced by the processing of a functionally distinct dimension (e.g. identity). While neuropsychological research (e.g. Garrido & et al., 2009; Hailstone, Crutch, Vestergaard, Patterson, & Warren, 2010; Humphreys, Donnelly, & Riddoch, 1993; Winston et al., 2004) as well as some behavioural studies (e.g. Campbell, Brooks, deHaan, & Roberts, 1996; Ellis, Young, & Flude, 1990; Young, McWeeny, Hay, & Ellis, 1986) support the notion of independent pathways in the processing of identity and affect, other behavioural studies (Ganel & Goshen-Gottstein, 2004; Schweinberger, Burton, & Kelly, 1999; Schweinberger & Soukup, 1998) as well as a neuroimaging study (Fox, Moon, Iaria, & Barton, 2009) suggest an *inter*dependence between variant and invariant aspects of the face, especially between identity and expression. Belin et al.'s model (2004) assumes independence of (lip) speech, identity and expression both at the uni-modal and bi-modal level. Experiment 2 addressed whether cross-modal adaptation is contingent upon the adaptor and test stimuli coming from the same, familiarised identity or whether these aftereffects are identity independent.

The present study examined some of the predictions of the 'auditory face' model (Belin et al., 2004). Using dynamic video stimuli, Experiment 1 tested whether adaptation to angry faces would shift perception of affectively ambiguous voice morphs towards fear and whether repeated exposure to fearful faces would shift perception in the opposite direction. Experiment 2 aimed to explore the relationship between identity and emotion at a supra-modal level of representation. If identity and affect are processed by dissociable pathways, it is likely that any cross-modal aftereffect should be robust to changes in the identity between face adaptor and vocal test morphs. Thus, we predicted that adaptation to one individual's face would elicit aftereffects in the perception of a different individual's voice.

2. Methods

2.1. Participants

Student volunteers from Bangor University participated in return for course credit. Participants had normal or corrected-to-normal vision and normal hearing, were Caucasian due to known effects of race on face perception (e.g. Lindsay, Jack, & Christian, 1991) and had no psychiatric or neurological problems. Twenty-five volunteers contributed to Experiment 1 (19 female, mean age = 23.4, standard deviation (SD) = 6.1) and twenty-three participants contributed to Experiment 2 (15 female, mean age = 20.7 SD = 4.09). None of the participants took part in more than one reported study and none were personally familiar with any of the age-matched individuals displayed in the stimulus material.

2.2. Stimuli

A pilot study using static images of emotional faces resulted in no cross-modal aftereffects (see supplementary online material (SOM) for details). We therefore created full high-definition videos of twelve individuals displaying seven expressions (6 basic emotions and neutral) in a professional sound studio with controlled lighting. We recorded videos and high quality audio simultaneously. Volunteers were instructed to use the vowel "ah" for each expression.