What makes a voice ‘mine’? Investigating the roles of ownership, choice, and agency in the processing of self-associated voices.

Bryony Payne

Thesis submitted for the degree of Doctor of Philosophy
University College London

Declaration of Authorship

'I, Bryony Payne, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.'

Signed: B Payne

Date: 16.12.21

Abstract

This thesis explores what it is that makes a voice 'mine'. It examines the influence of ownership over a new voice, personal choice in selecting it, and agency in using it on whether an unfamiliar voice is processed as 'self'. Experiment 1 shows that an unfamiliar voice can become associated with the self through ownership and subsequently prioritised in perception as a self-relevant stimulus. Experiment 2 shows that this perceptual prioritisation – and so self-bias – is not increased if the new self-voice is more representative of the self. However, Experiment 3 shows that bias *is* influenced by personally choosing the self-voice. Experiments 4 and 5 show that people also experience a greater sense of agency over a new voice that they own relative to a voice that is other-owned, which supports the finding that a new voice has become part of the self-concept. Experiment 6 demonstrates that self-bias for the *true* self-voice remains greater than bias for a new voice but, conversely, sense of agency is similar for the two. Experiment 7 shows that the greater bias for the true self-voice is diminished when that voice is presented as being owned by an 'other'. Moreover, under these conditions, participants retain a greater sense of agency over a new self-voice relative to the true self-voice. Experiments 8 and 9 then show that the bias for a new self-voice does not extend to better memory for information expressed in that new voice. Finally, Experiment 10 shows that using a new self-voice via text-to-speech technology does not influence the degree to which that voice is perceptually prioritised, nor the sense of agency experienced over it. Overall, this thesis demonstrates that it is the knowledge that a new voice is self-owned and so 'mine' that quickly and pervasively shapes perceptual processing and sense of agency.

Impact Statement

This thesis provides the first behavioural investigation into the possibility of incorporating a new voice into the self-concept. The implications of this work are both academic and applied. First, I demonstrate that a complex biological stimulus that objectively belongs to another person (i.e., their voice) can immediately acquire self-relevance through a sense of ownership over that voice. In becoming self-relevant, the way in which that voice is processed is fundamentally altered; it is prioritised in perception over the voices of others. Moreover, having ownership over a voice affects the sense of agency people feel over it; when a self-owned voice is heard as the outcome to self-action, the time between that action and its outcome is more compressed than when the outcome is not self-relevant. This influence of ownership on processing occurs robustly, despite the ownership being arbitrary and transient.

This contributes to our understanding of the pervasive influence of 'the self' – and what is deemed relevant to the self through ownership – on our wider processing. Particularly, it furthers our understanding of the factors that allow stimuli to be dynamically marked as self-relevant or not; ownership alone is sufficient for the integration of a new voice into the self-concept. However, I also show that having personal choice over which voice to select or reject as a new self-voice can also modulate processing. Voices that are rejected as 'self' are comparatively deprioritised and deemed less relevant. The studies within this thesis additionally show that the similarity of physical properties between a new voice and its owner does not make a voice any more self-relevant, nor does the ability to have control and agency in using it.

This insight into the factors that enable a new voice to be perceived as 'self' can inform the latest developments in voice synthesis technologies and, particularly, benefit the design and selection of individuated voices used within assistive communication devices. People who lose the ability to speak with their biological voice (e.g., people with motor neurone disease / ALS) can come to rely on such devices as a new means of self-expression. However, they often report that in losing their self-voice, they feel a loss in their sense of self. This research suggests that by giving people ownership of a new bespoke voice, they may gain a voice that can become uniquely 'theirs'. This ownership may allow that voice to become part of the self-concept and be processed as a self-voice. Moreover, the opportunity to choose or reject which voice to own may give people agency in choosing ‘what is self’ and what is not. This choice can alter the way in which those voices are subsequently perceived. Overall, this body of work significantly furthers our understanding of the flexibility of the self-concept, the influence of self-relevance on processing and may benefit people who need to incorporate a new auditory identity into their concept of self.

Contents

Tables

Figures

1 General Introduction

“Being human means being conscious of having a self and the nature of self is central to what it means to be human”

- (Oyserman, 2001, p. 499).

“*The human voice is an embodiment of self…contributing to the mutual exchange of self, consciousness, inner life, and personhood*.”

- (Sidtis & Kreiman, 2012, p. 146)

1.1 The self-concept

The self is a multifaceted construct, and one that is central to our cognition. Each person’s concept of self is comprised of a wealth of information: information related to our bodies, our subjective personality traits, values and beliefs, competence and successes, even the social roles we play and the possessions we own (Mittal, 2006; Belk, 1988). Thus, the self is not only defined by the *physical* self – what we look like, sound like, or our bodily motion – but rather by an array of self-knowledge that is built, maintained, and updated by interaction with our external environment.

Necessarily then, the boundaries of the self are fluid (Heersmink, 2020) such that we are able to accommodate new influences and information within the self-concept. This is essential if we are to be able to shift our understanding and recognition of self over time, perhaps as our appearance changes as we age, or we become more or less able-bodied, as we update our socio-political beliefs or the in-groups we feel we belong to. The self-concept can – and must – accommodate shifts in biological, artefactual (i.e., possessions), and sociocultural structures (Heersmink, 2020). To this end, we are all “soft selves, wide open to new forms of hybrid cognitive and physical being” (Clark, 2007, pg.279).

The flexibility of self-representation is evident from studies such as the Rubber Hand Illusion (Botvinick & Cohen, 1998; see also the enfacement illusion, Sforza, Bufalari, Haggard, & Aglioti, 2010; Tsakiris, 2008). In such studies, the participant positions their arm and hand out on a surface in front of them and these are obscured from view. A rubber hand is then placed on the same surface such that it is visible and positioned congruently with the body. Both the rubber hand and the biological hand are then stroked simultaneously by an experimenter so that the participant can see the tactile movement on the rubber hand only. The synchronicity between the visual and tactile cues is sufficient for participants to then perceive the rubber hand as part of their bodily self; participants feel as if it is their own hand. Indeed, if they perceive a threat or potentially painful stimulation to the rubber hand, they will react (withdraw) as if the incoming danger could affect them.

This effect relies on the fact that the sensory information (i.e., visual cues and tactile stimulation) occurs within the space immediately surrounding the body. The self-concept includes representation of this space, called the peripersonal space (Brozzoli, Makin, Cardinali, Holmes, and Farnè, 2011). This is the space within which we can reach, and be reached by, external entities (Rabellino, Frewen, McKinnon, & Lanius, 2020). Monitoring this space, via incoming multisensory stimuli occurring on or close to the body, helps us to avoid colliding with other entities (de Haan et al., 2016) and guides the grasping of objects (Brozzoli, Claudio, Pavani, et al., 2009). Importantly, multisensory integration is enhanced within this space such that the visual feedback of tactile stimulation on the rubber hand is sufficient to generate an illusory feeling of self-ownership over the hand.

Relatedly, the remit of this peripersonal space can be permanently extended by the frequent use of tools or external objects (Serino et al., 2007; Biggio, Bisio, Avanzino, Ruggeri, & Bove, 2020). For instance, tennis players' representation of this self-space is extended – relative to non-tennis players – to include their racket (Biggio et al., 2017). This extension is evident in behavioral measures and is also represented neurally, with structural and functional changes to the brain (Fourkas et al., 2008; Di et al., 2012). Critically though, this self-space only extends to objects that we have ownership over (Patané, Brozzoli, Koun, Frassinetti, & Farnè, 2021).

Indeed, anything that we come to self-own (Belk, 1988; Mittal, 2006;) or have agency in controlling becomes – at least transiently – incorporated into part of the self. This need not be a physical extension of the self as in the peripersonal space but, more simply, anything that we have a feeling of "mineness" over (Gallager, 2000), i.e. *"This* *is 'my' voice", "That is 'my' theory", or "I 'own' that car".* Thus, we have a sense of ownership in relation to stimuli that comprise the self-concept. Further, we are typically able to act on, and exert control over, the things that we own. According to Gallager (2000) a sense of agency is also central to the self-concept. The sense of agency refers to the feeling of being an agent that exerts influence on the external environment. For instance, the judgement that "*I moved my arm*", "*I swung my racket,*". In other words, it is the feeling of authorship over the outcomes we have – or believe we have had – control over (see Braun, Debener, Spychala et al., 2018 for review). Sense of agency helps us to distinguish which sensory stimuli we ourselves have caused from those that are externally generated, and thus is essential in differentiating self from other. Therefore, the sense of ownership and sense of agency are key mechanisms through which we both construct and experience what is self and what is not self (Gallager, 2000; Haggard et al., 2002; Braun, Debener & Spychala et al., 2018). These key concepts are discussed further in Chapters 2 (ownership) and 3 (agency).

1.2 Biases for 'self'

Although what is, and what is not, 'self' can be continually and flexibly constructed, the self-concept – as a cognitive structure – functions as a stable anchor point (Oyserman, 2001). Further it is an anchor point that influences how we perceive and act on the external world. Thus, what constitutes the self may be dynamically constructed, but whatever content currently comprises the self can influence – and bias – on-going information processing. Indeed, the self-concept "serves to bolster the stability of its components via enhanced stimulus processing" (Golubickis, et al., 2019, pg. 34). For instance, previous studies have shown that cognitive processes underlying *attention* (Shapiro et al., 1997; Tacikowski & Nowicka, 2010; Truong & Todd, 2017), *perception* (Sui et al., 2012; Constable, Welsh, Huffman, & Pratt, 2019), *judgement* (Mezulis, Abramson, Hyde, & Hankin, 2004; Hughes and Harrison, 2003), *decision making* (Li, Zhu, Ding, Ren, & Luo, 2019), and *memory* (Cunningham et al., 2008; Conway & Pleydell-Pearce, 2000) are all influenced by the self-relevance of information. This manifests as quicker recognition (Devue & Brédart, 2011; Ma & Han, 2010), faster rate of learning (Veldhuis, 2019), better memory (Brédart, 2016; Cunningham, Turk, McDonald, & Macrae, 2008; Kesebir & Oishi, 2010) and more positive judgements

(Mochon, Norton, & Ariely, 2012; Repp & Knoblich, 2004) for self-relevant information compared to information related to others.

1.2.1 Ownership

Importantly though, what is deemed self-relevant will change from moment to moment according to the currently salient content comprising the self-concept (Oyserman, Elmore, & Smith 2012). As such, new external stimuli can become relevant to the self-concept, particularly through ownership. Taking ownership of something affords it status as a self-relevant stimulus and it is, thereafter, processed preferentially with bias. For example, if people are told they have a theory about an abstract concept while another person has an opposing theory, people will be more likely to believe and endorse their 'own' theory, simply because it belongs to them (Gregg, Mahadevan, and Sedikides, 2017). People also judge items they own more positively (Huang, Wang, Shi, 2009) and value them more highly than identical objects owned by others (Dommer and Swaminathan, 2013). Interestingly, the value increases according to the duration of the ownership (Strahilevitz & Loewenstein, 1998).

This bias also affects perceptual processing. Specifically, people are faster at identifying objects that they own and also in making judgements about whether they are self-owned or not. For instance, Falbén, et al. (2019) assigned participants either a pen or a pencil to own, while a 'friend' was assigned the other object. At test, participants were shown an object onscreen and were asked to make speeded judgements pertaining to ownership and object identity. In one block, participants judged whether the object belonged to them or to the friend while, in another block, they judged whether the object was a pen or a pencil. In both instances, the object that participants had been given ownership over was reacted to faster than the other-owned object. This effect is known as the self-prioritisation effect (Sui et al., 2012), which is expanded upon in Chapter 2.

Furthermore, the bias for self-owned items is evident in memory. Specifically, Cunningham et al. (2008, see also Van den Bos et al., 2010) gave people ownership of various food items by placing them into a shopping basket assigned to them. Each item was represented pictorially so ownership was imaginary, arbitrary and temporary.

Despite this, in a surprise recall test, participants showed a memory advantage for the items they owned relative to the items that had been put into another’s basket and were therefore other-owned.

1.2.2 Choice

Interestingly, this memory advantage is increased if the objects are not only self-owned but personally chosen (Cunningham, Brady-Van den Bos & Turk, 2011). That is, items that people personally choose to take ownership of are better remembered than items they are randomly assigned ownership of. On the one hand, chosen items are likely to be more congruent with information currently comprising the self-concept, i.e. an item will be chosen because it is in line with personal tastes and preferences. However, previous studies have shown that having a choice is, itself, inherently rewarding (Leotti & Delgado, 2015). Further, the relationship between self and choice is bidirectional; people make choices in line with their personal preferences and, in turn, their choices positively bias their preferences (Alós-Ferrer & Shi, 2015; Ariely and Norton, 2008). Thus, between two very similar items, people will prefer the object that they have chosen (Huang et al., 2009). The act of choosing – perhaps as a means of exerting ‘self’ within the external world – induces bias for the outcomes of those self-made choices. As Shang, Tao, & Wang (2016, p.2) write, ‘“my choice” is also a part of the self-concept’ and so that choice is intrinsically self-relevant and attributed bias.

1.2.3 Agency

The action of taking ownership and the action of choosing, can dynamically determine what is ‘self’ and what is self-relevant from what is not. Relatedly, the way in which the self physically interacts with external objects shapes how they are subsequently processed. Truong, Chapman, Chisholm, Enns, & Handy (2016) demonstrated that the saliency of an item is increased if it is physically moved closer to the self. Further, Huffman & Brockmole (2020), demonstrated that people’s attention was biased towards stimuli that they had agency over and were able to control. Participants were presented with moving visual stimuli (i.e., circles onscreen), a selection of which they had control over. In different trials, they had a different degree of control. The task required participants to make a key press when they perceived a change in the visual stimuli and the authors found that participants’ reaction times were significantly quicker when it was the stimuli under their control that changed. Furthermore, the more agency participants had in controlling the stimulus, the more biased they were towards it in terms of attentional selection, as measured by decreasing reaction times. Functionally, it is understandable that people are biased towards things that they have agency over. Such a bias may help people to monitor the outcomes they have had agency in causing, and to ensure the outcome is what they intended (Hauf, Elsner, & Aschersleben, 2004). Relatedly, the entities people have agency in controlling tend also to be the entities that are self-owned (i.e., our bodies, our voice, our material possessions). Thus, a sense of ownership and sense of agency, though distinct, often coincide and are central to the self-concept.

Overall then, the self-concept biases our cognitive processes towards what is self-relevant. What is self-relevant can be dynamically extended by, and experienced through, what the self has ownership, choice, and agency over. Thus, as Oyserman (2001, pg. 500) writes:

> *"The self-system is both an array of self-relevant knowledge, the tool we use to make sense of our experiences, and the processes that construct, defend and maintain this knowledge."*

1.3 Neural correlates of self-bias

The neural instantiation of self, and the breadth of self-relevant processing has long been under investigation and remains difficult to define. Previous studies have shown that the processes underlying self-bias are at least partially distinct from the familiarity of the stimulus (Sui et al., 2012) or its reward value (e.g., Yankouskaya et al., 2017) but whether there are neural processes specific to self-processing remains in question. Behaviourally, the biases for the self appear distinct across cognitive domains (i.e., memory, attention, perception). That is, the behavioral measures that are elicited across tasks are not typically correlated (Nijhof, Shapiro, Catmur, & Bird, 2020), which suggests no common, domain-general bias.

Notwithstanding, it is widely established in neuroimaging studies that self-referential processing is associated with the medial prefrontal cortex (mPFC), an area linked to self-representation (Kelley et al., 2002; Mitchell, Macrae, & Banaji, 2006). Specifically, ventral regions of the mPFC are relatively more active for self-relevant processing compared to other-related processing (Benoit, Gilbert, Volle, & Burgess, 2010). Conversely, more dorsal regions of the mPFC are associated with other-relevant processing in comparison to self-relevant processing (Qin & Northoff, 2011; van der Meer et al., 2010). However, that is not to say that these areas are self- or other-specific. Indeed, making judgements about other people that we perceive to be similar to ourselves activates the ventral regions significantly more than when we think about dissimilar others (Denny, Kober, Wager, & Ochsner, 2012). This suggests that the greater activation of more ventral regions of the mPFC is a function of a stimulus' self-relevance.

Sui, Rothstein, and Humphreys (2013) have also shown that the facilitation of newly self-relevant stimuli in perception, is reflected in increased functional connectivity between the ventral medial prefrontal cortex (vmPFC) and the posterior superior temporal sulcus (pSTS), an area linked to social attention (Allison, Puce, McCarthy, 2000; DiQuattro & Geng, 2011). Indeed, the coupling strength between vmPFC and pSTS is correlated with the degree to which self-relevant stimuli are facilitated in perception. In linking a neutral stimulus to self-representation, social salience for that stimulus is increased because its self-relevance is increased. As a stimulus' self-relevance increases, so too does activity in vmPFC (Sui, Rothstein, and Humphreys, 2013; Humphreys & Sui, 2016). Thus, an external stimulus that becomes associated with the self via ownership does become part of the self-concept, at least in as far as it is processed as self-relevant.

However, this coupling may be part of a much larger network. Given the multifaceted nature of the self, its neural underpinnings are likely to be widely distributed (Gusnard, Akbudak, Shulman, Raichle, 2001). Further, any such network must, necessarily, connect social, self-referential and high-level cognitive control processes. Currently, the most well-established network proposed to underlie the self and self-processing is the default mode network, which includes the medial prefrontal cortex (mPFC), the anterior (ACC) and posterior cingulate cortex (PCC), lateral and medial temporal lobes, and posterior inferior parietal lobule (pIPL). Within this network, the mPFC is posited to be a 'social information hub' (Meyer & Lieberman, 2018) that is repeatedly coupled with other sources of socially-driven, self-relevant information. The vmPFC is thus an organiser of self-relevant knowledge, able to facilitate and prime access to such knowledge when it is needed (Meyer & Lieberman, 2018). According to Meyer & Lieberman (2018) it is essential that we are able to access self-relevant information at short notice to be successful in meeting social demands. Indeed, we can only successfully navigate social interactions if we understand the self and, further, understand the self in a complex social network (Meyer & Lieberman, 2018).

1.4 The social construction of self

It is broadly agreed that self-processing is socially motivated (Yin, Bi, Chen, Egner, 2012) because, inherently, the concept of what is, and what is not, 'self' is socially constructed (Forgas and Williams, 2002; Northoff, 2013).

In infancy, the boundaries of self and other are constructed in terms of the physical self; understanding that the self-body is separate and different to another's body (Oyserman, 2001). Thereafter, as we come to interact with the environment, we continually redefine and renegotiate self and other according to social structures. For instance, a person may be perceived as being more distinct from the self – more 'other' – if there are no shared values, beliefs or experiences (Oyserman, Elmore, & Smith, 2012). There is a wealth of evidence that has demonstrated that people are, in fact, biased towards preferring people they perceive to be similar to the self (Gerson, & Bekkering, & Hunnius, 2017; Mahajan & Wynn, 2012; Jones, Pelham, Carvallo & Mirenberg, 2004). Similar others are deemed to be more self-relevant than dissimilar others and so the bias for the self extends to them. In this way, people have a better memory for objects and information associated with people who are similar to them (Allan, Morson, Dixon, Martin, & Cunningham, 2017). Further, people find similar others (Jones, Pelham, Carvallo, & Mirenberg, 2004) and their voices (Peng, Wang, Meng, Liu, Hu, 2019) more attractive.

Importantly then, people's values, beliefs and experience – indeed peoples' whole selves – can only be shared through interaction with others. One of the primary means by which this is achieved is through the use of the voice. The voice is an essential conduit through which people represent themselves. Relatedly, it is from listening to other people's voice that we infer their 'self'. In this way, a person's voice is central to their self-concept – to its construction, maintenance and expression. The importance of owning a self-voice and what this means for investigating the self and self-bias is explored further below.

1.5 The self-voice

Each person has a unique voice pattern that arises from both physical differences in their vocal apparatus and also from the flexibility with which it can be used (Scott & McGettigan, 2015). Thus, each person's voice is idiosyncratically marked by aspects of self and, through the use of the voice, a person can share their 'self' with others.

A substantial body of work has now demonstrated that a people can perceive a person's sex (Puts, Gaulin, & Verdolini, 2006), age (Smith & Patterson, 2005), body-size (Pisanki, 2014), physical strength (Sell et al., 2010) and health (Pribuisiene, Uloza, Kupcinskas, & Jonaitis, 2006) from an individual's voice (Belin, Fecteau & Bédard, 2004; Kreiman & Sidtis, 2011; Mathias & von Kriegstein, 2014). Further, people can form an impression of a persons' level of education, occupation, regional origin, ethnicity, and social status (for review, see Kreiman & Sidtis, 2011). Thus, a wealth of information that has contributed to the construction of the self-concept can be manifested in the voice – the 'embodiment of self' (Sidtis & Kreiman, 2012, p.146).

It is understandable then how the voice has previously been referred to as an 'auditory face' (Belin, Bestelmeyer, Latinus, & Watson, 2011) given the similarity of person-related information that may be perceived via face and voice. Indeed, people may also judge how extroverted (Scherer, 1978) or neurotic (Hu, Wang, Short, & Fu, 2012) someone is from the sound of their voice. Moreover, people can – and do – perceive social traits such as whether that person is trustworthy, attractive, intelligent, or dominant from very brief exposure to their voice (i.e., on hearing "Hello") (McAleer, Todorov & Belin, 2014). Importantly, these judgements – often first impressions – can then bias subsequent decision-making and attitudes towards that person. For instance, people perceive whether a speaker sounds confident or not after as little as 200ms (Jiang & Pell, 2015) and, if a speaker sounds confident, they are judged to be more believable (Jiang & Pell, 2017). Moreover, by association, those speakers are also perceived to be more intelligent and hold higher social status (Jiang, Gossack-

Keenan, & Pell, 2019). Unsurprisingly then, judgements about a person's voice can affect their chances of attaining employment (Schroeder & Epley, 2015) and, more broadly, can influence election outcomes (Banai, Laussten, Banai & Bovan, 2017). Thus, implicitly, judgements about a person's voice – as a proxy measure of self – affect how successfully they can navigate the social world. It is critical that the voice portrays the self favourably or, at least, as the speaker would want to portray themselves.

Although the self is represented by multiple modalities (self-face, bodily motion), the self-voice is the tool with which people have the most agency in controlling *how* it represents them. Unlike the self-face, which represents someone's identity just by existing and being seen, the self-voice cannot exist passively. A voice always requires intentional action to exist (Scott and McGettigan, 2015) which, in part, is why it is so representative of that speaker: the existence of the self-voice demands purposeful behaviour to express oneself. A speaker can extensively and flexibility modify their voice to fulfill diverse communicative goals and this can be done dynamically, situation to situation (McGettigan et al., 2013; McGettigan & Scott, 2014). Thus, we can "adopt a variety of 'selves'" (McGettigan, 2015, pg. 2) to signal aspects of our identity and represent the self appropriately, according to changes in social demands and the acoustic environment.

For instance, even if the semantic content is equivalent, people's use of their voice will differ according to whether they are speaking at a wedding or a funeral, talking to a grandparent, an infant, or a police officer (Lavan, Burton, Scott & McGettigan, 2018). Indeed, the flexibility of using the self-voice extends to nuanced and modified speaking styles dictated by whether a speaker is trying to convince, commiserate, motivate or defend their interlocutor, for instance. Guldner, Nees, & McGettigan, (2020) have recently shown that people are successful in modulating their voice to more strongly express certain social traits; listeners are sensitive to these modulations and perceive the voice as intended, for instance, as a voice that sounds confident or hostile. Thus, the voice is a tool with which people can navigate complex social contexts, presenting a self that is congruent with – and optimised to – the audience, the acoustic environment and their communicative goals.

The fact that voice is a means through which the self-concept is negotiated and expressed – particularly in relation to others – is illustrated through the phenomenon of vocal convergence. Often without being aware, people frequently converge their acoustic-phonetic pronunciation to merge with their interlocutor (Pardo, 2006). Previous studies have shown this convergence reflects, and helps to navigate, the social distance between people (Giles, Coupland, & Coupland, 1991) and may also increase affiliation between speakers (Pardo, Gibbons, Suppes, & Krauss, 2012). Indeed, the degree of phonetic convergence has been shown to be related to speakers' feelings of social closeness (Pardo et al., 2012). Similar to how we can consciously modify our speaking style to fit a certain situation then, convergence is another illustration of how the self we present through voice is co-constructed according to the other.

Overall, it is clear the self-voice is intrinsically tied to the self-concept as a means of constructing, maintaining and expressing what is self in relation to others. Indeed, the "voice is revelatory of 'self', mental states, and consciousness, and reflects both the speaker and the context in which the voice is produced," (Sidtis and Kreiman, 2012, pg. 4). The importance of owning a voice and, further, of having agency in using that voice, becomes particularly apparent when the impact of vocal impairment or voice loss is considered.

1.5.1 The importance of having a self-voice

Vocal impairments or a reduction in vocal function can critically impact people physically, socially, and emotionally (Porcaro et al., 2021) For instance, vocal impairments arising from a reduction in lung function (i.e., after spinal cord injury, Nygren-Bonnier et al., 2011) can reduce someone's ability to speak loudly. Thus, effective communication is jeopardized in noisy environments, when speaking across multiple talkers or when projecting the voice to give a presentation (Nygren-Bonnier et al., 2011; Ward, Jarman, Cornwell, & Amsters, 2016). This increases the vocal effort required, which can lead to frustration and anxiety at having to repeat oneself to be heard or understood (Johns, Arviso, & Ramadan, 2011). Similarly, people judge a quieter, aging voice to be weaker and this judgement can affect how the whole "self" is perceived. Montepare, Kempler, & McLaughlin-Volpe (2014) found that more elderly speakers were judged to be less powerful and less engaged because of their aging voice quality.

The judgement of someone's self according to their voice occurs even amongst children. It has previously been shown that dysphonic children are (mis)judged by non-dysphonic children to have relatively less positive qualities. For instance, they were judged to be significantly less *kind, happy, honest, good*, or *clean* relative to non-dysphonic children (Lass, Ruscello, Stout, & Hoffmann, 1991). Perhaps unsurprisingly then, dysphonic children can report feeling angry, sad, or embarrassed about their voice which, without intervention, can lead to a loss of self-esteem and increasing social exclusion (Connor et al., 2006). In adulthood too, a person speaking with a voice impairment may be (mis)perceived as less credible (Schroeder, Rembrant, May, & Freeman, 2020), less intelligent, or less socially desirable (Allard & Williams, 2008). These perceptions can negatively bias how successful that speaker is in attaining employment (Schroeder, et al., 2020) or the social activities they participate in.

This wealth of research shows that the self-voice is critical to the social construction and expression of self. Judgements based on a person's voice act as a proxy measure of self: "As another persons' mind cannot be experienced directly, its quality must be inferred from indirect cues", (Schroeder, Kardas & Epley, 2017, p.1760). Thus, a loss of voice, or a loss of agency in using that voice, can be a loss of self. Indeed, people who have lost their ability to use their natural voice have previously reported feeling "semi-detached" (Cave & Bloch, 2020) from conversations. In losing a flexible means of self-representation, people lose a conduit through which they can actively place themselves within, and contribute to, community and social relations.

1.5.2 Integrating a new auditory identity

Many people with a voice impairment or those who lose their ability to speak with their natural voice come to rely on an alternative means of communication. This includes children and adults with diverse diagnoses, including cerebral palsy, locked-in syndrome, autistic spectrum disorders, and motor neurone disease (MND) (Langarika-Rocafort, Mondragon, & Extebarrieta, 2021). The case of motor neurone disease is particularly poignant because people who are diagnosed with MND will, typically, have used their natural self-voice for the majority of their lives. After symptom onset, however, 80–95% of people with MND lose the ability to use their own voice to communicate their daily needs (Beukelman, Fager, & Nordness, 2011). Therefore, they may choose to rely on an alternative means of communication (Cave & Bloch, 2020).

Augmentative and alternative communication (AAC) devices are varied in their function, method, and degree of sophistication. With recent improvements in speech technology, however, high-tech AAC devices using synthesized or digitized voices have become increasingly available (Banda & Alzrayer, 2018; see Langarika-Rocafort, Mondragon, & Extebarrieta, 2021 for review). These devices may either use a synthesised voice or, with specific companies such as CereProc and VocalID, a bespoke digitised version of a person's natural voice if it has been "banked" or preserved before loss (Pullin, Treviranus, Patel, & Higginbotham, 2017). Indeed, the option of voice banking for people with MND – that is, to pre-record their own voice for future use within an AAC device – is experienced almost unanimously as being for the sake of 'preserving identity' (Cave & Bloch, 2020). For instance, people report that the opportunity to keep their natural voice can help to *"*save a piece of themselves*"* (Cave & Bloch, 2020, p.123).

However, more typically, and in instances in which voice banking was never possible, AAC devices use non-bespoke and somewhat generic synthesized voices. Although it has recently been recognised that personalised voices are important for individuating a personal identity and for social integration (Mills, Bunnell, & Patel, 2014; Pullin, Treviranus, Patel, & Higginbotham, 2017), such options have only just started to emerge. Despite the expansion of digital voices in the tech industry (i.e., Alexa, Siri), more bespoke options for use with AAC devices have not made as much headway. The progress that has been made has been predominantly motivated by the intuitive, moral case for providing individuated and personal means of self-expression.

However, the importance of having a voice that represents one's own auditory identity has not been examined scientifically, nor has the possibility of incorporating a new auditory identity into the self-concept. These questions underpin the research examined with this thesis, which is expanded below.

1.6 The current thesis

The previous sections of this chapter have illustrated the extent to which the self is a socially developed cognitive structure that biases our cognition. Across domains, including attention, perception, and memory, it is clear that incoming information is processed preferentially and with an advantage if that information is self-relevant. However, what is self-relevant will shift from moment to moment according to the currently saliently content comprising the self-concept. Indeed, the self-concept is, itself, flexible; what is 'self' may be dynamically constructed and extended through interaction with the external environment and with other 'selves'.

A person's voice is central to the self-concept as it is a conduit through which the self can be socially constructed and expressed to others. Loss of the self-voice may negatively affect people, leading to social detachment and a loss in the sense of self. Although people may come to rely on a new, externally generated voice as an alternative means of self-expression, it remains in question whether this new voice can become integrated into the self-concept and be processed as self-relevant. This clinically pertinent question may inform the design and uptake of new voices. Further, and perhaps more importantly, it also forces an examination of what it means to have a voice that is one's own.

This thesis therefore explores what makes a voice part of the self, focusing on the factors that may influence whether a voice is processed as self-relevant or not. Further, it investigates how a voice becoming self-relevant affects our experience of that voice. This is examined particularly in terms of how we perceptually process it and experience a sense of agency over it. This is achieved through a series of behavioural experiments, outlined below.

Chapter 2 explores whether a new externally-generated voice can become associated with the self-concept through ownership and subsequently processed as self-relevant. Self-relevant information biases perceptual processing and typically elicits faster and more accurate perceptual judgements. This is known as the self-prioritisation effect (Sui et al., 2012). Experiment 1 thus explores whether a new self-owned voice is perceptually processed as self-relevant according to whether it is judged more quickly and accurately than other-owned voices. Experiment 2 furthers this examination by testing whether a new self-owned voice is attributed greater self-bias (i.e., is deemed more self-relevant) if it is more representative of the self, specifically in terms of matching the participant's gender-identity. Finally, Experiment 3 asks whether this bias is increased by personally choosing a new voice to own, relative to its being arbitrarily assigned. All three experiments use a perceptual matching paradigm, which is detailed further within the chapter.

Chapter 3 then explores whether a voice that has become self-relevant through ownership influences the sense of agency people experience over it. Sense of agency is measured implicitly using a temporal binding paradigm which is detailed further within the chapter. I first report a pilot study which was conducted to optimize the adaption of a temporal binding paradigm for online use with auditory stimuli. This pilot study has implications for the experimental design and interpretation of data in Experiments 4 and 5. Experiment 4 then examines the sense of agency over a new self-owned voice relative to an other-owned voice. Experiment 5 furthers this by exploring how postdictive cues regarding whether a voice is self-owned or other-owned may influence participants' sense of agency.

Chapter 4 asks how both perceptual processing and sense of agency is influenced by the self-relevance of a voice. Specifically, in Experiment 6 I assess people's experience of a new self-owned voice as compared to their own, true self-voice. The true self-voice is a highly salient, self-relevant stimulus that inherently signals the participant's unique identity. Measuring how the true self-voice is perceptually prioritised and, further, the degree to which participants experience a sense of agency over it, can inform understanding of how far a new self-owned stimulus has been processed as self-relevant by comparison. Thus, this experiment explores the degree to which an other-generated voice can be processed as 'self' in comparison to the true self-voice.

Experiment 7 further manipulates what is ‘self’ and what is ‘other’. Here participants are given ownership of an other’s voice to own as a new self-voice while, conversely, being presented their true self-voice as being other-owned. Such a manipulation elucidates the extent to which self-bias is flexibly applied according to the context and, further, how participants’ sense of agency is influenced dynamically by what they believe is most self-relevant.

The self-concept, and information that is deemed relevant to it, has been shown to bias multiple different domains. Chapter 5 therefore extends the examination from perception to memory. In Chapter 5, Experiment 8 asks whether, through stronger encoding for self-relevant information, people demonstrate better memory for items said aloud in a new self-owned voice relative to an other-owned voice. Experiment 9 furthers this examination into memory by asking whether source memory specifically (i.e. memory for whether it is self-owned or other-owned) is enhanced.

The final experimental chapter examines perception and agency over a self-owned voice when it is a tool that can be used to represent the self. Voices are fundamentally social stimuli, and their importance to the self may be underpinned by the extent to which they can be used to express the self and achieve communicative goals. Chapter 6 therefore asks whether a new self-owned voice becomes more self-relevant if it has been used socially as a means of self-expression. Thus, in Experiment 10, I detail work that used state of the art text-to-speech technology to allow participants to use new self-owned voices in a real social interaction with another participant. The influence of this social use of a voice is assessed according to whether people perceptually prioritise a self-owned voice more, and/or feel a greater sense of agency over it, if it has been used to represent the self.

Finally, Chapter 7 offers a General Discussion in which the results from the ten experiments are considered in relation to each other and discussed in tandem. I also discuss the implications of this body of research, its limitations, and the future work that could further inform this essential question of what it is that makes a voice ‘mine’.

2 Perceptual prioritisation of a new self-voice[1]

Abstract

Within Chapter 2, three experiments begin to explore what makes a voice be deemed 'self', and whether an unfamiliar, externally-generated voice can become incorporated into the self-concept. Information that is relevant to the self biases processing such that this information is perceptually prioritised over information related to others. If the self can be extended to incorporate new information – such as a new voice – that voice should be perceptually prioritised as self-relevant. In Experiment 1, participants learnt new associations between three unfamiliar voices and the three identities they belonged to (self, friend, other). In a task, participants then made speeded judgements of whether voice-identity pairs were correctly matched, or not (Sui et al., 2012). Results showed faster and more accurate responses to the new self-owned voice relative to either the friend-voice or other-voice, showing a clear self-prioritisation effect. Two further experiments then measured whether prioritisation was enhanced if the self-owned voice was matched to the gender-identity of the participant (Experiment 2) or if the self-voice was personally chosen by the participant (Experiment 3). Gender-matching did not significantly influence bias. However, when participants chose their self-voice, the friend-voice became relatively *de*-prioritised compared with when the identities were randomly allocated. These findings contribute to our understanding of the flexibility of the self-concept and the factors that may influence how – and the extent to which – the voice of an 'other' may become self-relevant.

[1] The experiments described within this chapter have previously been published: Payne, B., Lavan, N., Knight, S. and McGettigan, C. (2021), Perceptual prioritization of self-associated voices. Br. J. Psychol., 112: 585-610. https://doi.org/10.1111/bjop.12479.

2.1 Introduction

The self-concept comprises ‘internal representations that shape perceptions of how the self is related to one’s surroundings and to other people,’ (Chiu, Ho, Tollenaar, 2021, p. 1). Given that the self-concept is flexible, the way in which the self is related to the environment will change; what external information is – and is not – relevant to the self will change.

According to Schäfer, Wentura, & Frings (2020), self-relevance acts as an associative glue, binding together disparate content to create a network of information which should be prioritised in processing. Indeed, across domains, our cognition is biased towards self-relevant information. In perception, this manifests as quicker and more accurate perceptual processing of information that is self-relevant compared to information that is not self-relevant. Thus, if the self is extended to incorporate new information, that information should be prioritised as self-relevant.

Sui, He, & Humphreys (2012) introduced a perceptual matching paradigm through which unfamiliar external stimuli can become associated with the self-concept and, therefore, be made self-relevant. Specifically, participants were asked to form associations between three social identities (i.e. self, friend, other), and three shapes (i.e. circle, triangle, square). For example, the participants were told: ‘the circle represents you, the square a friend, and the triangle, an other’. These three shape-identity associations were thus established as ‘matched’ pairings. In a subsequent test phase, participants were presented with mixed pairings of a shape and an identity label onscreen and were required to make speeded judgements of whether or not the two stimuli were correctly matched. Through this task, it was possible to measure the bias afforded to each stimulus by virtue of which social identity it had become associated with. The results showed that the shape that had been associated with the self was responded to more quickly and accurately relative to the shapes associated with the friend or other. Thus, there was evidence of a perceptual bias towards the shape associated with the self (Sui et al., 2012), an effect which became known as the self-prioritisation effect.

Importantly then, this perceptual matching paradigm enabled an examination of the bias for stimuli that had become self-relevant outside of how familiar they were. This is key because self-relevant stimuli such as the self-face or self-voice – which are typically attributed a processing advantage (Bortolon & Raffard, 2018; Keyes & Dlugokencka, 2014) – are also highly, personally familiar. Familiarity itself may confer its own advantage to attention and perceptual processing. Therefore, the perceptual matching paradigm demonstrated that bias exists specifically for the self, outside of familiarity. Further, the task also made it clear that unfamiliar stimuli can become self-relevant very quickly, simply by tagging them as being associated with the self. Thereafter, self-bias can be attributed flexibly to these stimuli that have become self-relevant.

Since its introduction, this paradigm has been used prolifically but has also been well interrogated. For instance, Sui et al. (2012) have demonstrated that this effect is not underpinned by the social identity labels used within the task. Specifically, the self-prioritisation effect is not reliant upon the choice of (pro)nouns used (i.e., you, friend, stranger/other), the word length of the label, its word frequency, or how concrete/abstract the concepts of self, friend, or other are (Sui et al., 2012). Schäfer, Wentura, & Frings (2017) additionally questioned whether the grammatical distinctiveness of ‘you’ as a pronoun relative to the nouns referencing ‘friend’ and ‘stranger’ could give rise to the effect due to its potentially greater salience. However, they too concluded that the self-prioritisation effect is ‘a genuine self-related process’ that was unaffected by the grammatical distinctiveness of the labels (Schäfer, Wentura, & Frings (2017: 400).

This examination has been extended even further by Woźniak & Knoblich (2019); they created a perceptual matching task without any labels at all. Specifically, they asked participants to form associations between three unfamiliar symbols and three unfamiliar faces. These three pairings were then verbally associated with either self, friend, or other, such that the presentation of a paired symbol and a face were, together, associated with an identity. At test, participants were presented with mixed pairings and were asked to make speeded judgements on whether they were a correctly matched pair or not. Again, participants were faster and more accurate at judging the two stimuli that were associated with self, relative to the pairs of stimuli associated with a friend or other, even in the absence of labels. Thus, the knowledge alone that stimuli were associated with the self, gave rise to a self-prioritisation effect.

Further studies have explored the extent to which the self-prioritisation effect can be dissociated from other factors that may facilitate processing. For instance, it has been demonstrated that the effect can be dissociated from the emotional valence of the stimuli associated with self, friend, or other (Schäfer, Wentura, & Frings, 2020, Stolte, Humphreys, Yankouskaya, & Sui, 2015). Relatedly, the prioritisation for self-relevant stimuli is at least partially distinct from the reward value of the stimuli. Wang, Qi, Li, & Jia (2021), asked participants to form associations between the self and a low-reward, a friend and a medium reward, and an other and a high-reward. At test, participants showed a significantly greater perceptual bias for the self, relative to the friend or the other even though the reward it was associated with was comparatively smaller. In line with the wider literature, the authors thus concluded that the self-prioritisation effect is differentiated from – and prevails over – reward-based processing. Indeed, it is largely established that although the neural processes underlying reward do overlap with self-processing, they are at least partially distinct from the mechanisms underlying self-prioritisation (Yankouskaya, Humphreys, Stolte, Moradi, & Sui, 2017; Greck, Rotte, Paus, Moritz, Thiemann, Proesch, Bruer, Moerth, Templemen, Bogerts, & Northoff, 2008; Northoff & Hayes, 2011, Sui & Humphreys, 2015).

Simultaneous to this extensive interrogation, the perceptual matching paradigm has been used to robustly replicate a self-prioritisation effect over an array of stimuli. For instance, beyond the original geometric shapes (Sui et al., 2012), other visual stimuli including objects (Falbén et al., 2019) and Gabor patches (Stein, Siebold, van Zoest, 2016) have been used in successful replications. Further, Schäfer, Wesslein, Spence, Wentura and Frings, (2016) asked participants to form associations between the three social identities and tactile stimuli or auditory tones. For instance, participants were asked to associate three neutrally-valenced pure tones with the self, friend, and other. Participants were then presented with an auditory tone and an identity label and asked to make speeded judgements on whether or not they were correctly matched. The results showed a comparable self-prioritisation effect for the auditory tone that had been associated with self, relative to the tones associated with others. This study therefore extended our understanding of the extent to which perception could be shaped according to self-relevance. Indeed, it demonstrates that external stimuli of all sensory modalities can quickly become self-relevant and prioritised perceptually.

This is further exemplified by Payne, Tsakiris, and Maister (2017), who asked participants to form new self-associations to three unfamiliar faces. Participants were specifically told that one face now belonged to them, one belonged to a friend, and another to a stranger. Thus, implicit within the instructions, participants were given ownership of a new self-face by being told it belonged to them. Given that a face intrinsically represents an identity and, moreover, a different identity to the self, it should be processed as other-associated. In spite of this, Payne et al. (2017) demonstrated that participants had faster reaction times and greater accuracy to the face newly assigned to self, relative to the other faces. The fact that an unfamiliar face was prioritised perceptually, reflects that it has been perceived as self-relevant and perceived as a temporary extension of self.

While it is agreed that a new stimulus can be quickly incorporated into the concept of the self, it remains unclear whether this extension alters representation of the physical self. Payne, Tsakiris, and Maister (2017) therefore ran a further study to assess whether or not participants' recognition of what was 'self' had changed to accommodate the new self-associated face. Specifically, they used a self-recognition task in which the participants saw images of their true self-face morphed with the newly self-associated face. The task comprised 50 different images such that each image represented a 2% incremental step in the morph between the two faces. At test, participants then rated how similar each image was to their real face. Importantly, this self-recognition task was completed both before and after the perceptual matching task such that the authors could track any change in self-recognition after a new, unfamiliar face had become self-relevant. However, the results showed that the percentage of morphing at which participants judged an image to be 50% of the real self-face and 50% of the new self-associated face did not differ between the pre- and post-tasks. That is, there was no detectable increase in perceived similarity between their own face and a newly self-associated face. This suggests that self-recognition had not been influenced by the integration of a new self-face into the self-concept. Thus, the self may be temporarily extended to encompass new stimuli conceptually through the perceptual matching paradigm, but this may not reflect changes to the representation of the physical self.

However, this result is at odds with other studies which typically show that self-representation can be modified by simple psychophysical manipulation. For instance, if a participant's face is touched simultaneously to them viewing synchronous touching of a partner's face, the boundary of self and other becomes blurred. Specifically, participants report that the partner's face is more similar to their own after synchronous touching, an effect that is known as the enfacement illusion (Sforza et al., 2010; Tajadura-Jiménez, Grehl, & Tsakiris, 2012; Paladino, Mazzurega, Pavani, & Schubert 2010). Indeed, as part of the enfacement effect, the point at which participants accept that a morphed image is in fact 'self' is typically reduced after the manipulation. For instance, Tajadura-Jiménez, Grehl, & Tsakiris (2012) demonstrated that, after the enfacement task, participants accepted an image of a morphed self-face as 'self' when it contained 5% more of the other's face than before the manipulation. This shows that participants perceived greater similarity between the self-face and other-face after manipulation. Further, this increased similarity rating was correlated with an increase in perceived empathy and attractiveness of the other's face (Sforza et al., 2010), as well as the social closeness felt between them (Paladino et al., 2010). This indicates that the enfacement illusion influenced not only the participants' body perception but also their social perception of what was 'other'. Thus, according to Sforza et al. (2010), the partner's face does become integrated into the representation of the participant's own face.

It is possible then that bodily self-representation is malleable but through mechanisms other than those engaged in the perceptual matching task. However, Sui et al. (2013) ran a study in which participants associated a new stimulus with the self through the perceptual matching paradigm while fMRI data was obtained. Behaviourally, participants demonstrated the typical self-prioritisation effect in which new external stimuli associated with the self were judged more quickly and accurately than stimuli newly associated to either a friend or an other. Neurally, the new self-associated stimulus elicited enhanced activity in the ventral medial prefrontal cortex (vmPFC), an area associated with self-representation (Northoff et al., 2006). Indeed, it is the vmPFC that is activated by highly self-relevant information such as a participant’s own name (Moran, Heatherton, & Kelley, 2009). Thus, its relatively greater activation for an unfamiliar stimulus newly associated with the self – in comparison to the equally unfamiliar stimuli associated with others – suggests that the self-stimulus had been, at least transiently, linked to self-representation (Sui et al., 2013).

Overall then, the current literature shows that external stimuli can be made self-relevant by becoming associated with the self. In becoming self-relevant, such stimuli are prioritised perceptually and responded to more quickly and accurately. Previous studies have shown that external stimuli can be encompassed within the concept of self. As far as it is currently possible to elucidate the neural underpinnings of self-processing and self-prioritisation, it remains unclear whether taking ownership of that stimulus is, alone, sufficient to fundamentally alter self-representation.

Thus far, Payne et al.’s study in which participants were given ownership over a new self-face provides the closest test case for which complex social stimuli may be accepted as part of the self-concept. However, whether a new voice can be integrated into the self-concept has never been explored. Voices are, like faces, highly social and identifying signals, yet they differ in that they require self-generated action in order to exist. It is only through the varied, idiosyncratic, and deliberate use of our voice that we can share it with others to express our self-identity. Relatedly then, a voice – in its very existence – is the inherent biological property of an individual and should therefore signal only that individual’s identity. To demonstrate that an unfamiliar voice, generated by another person, may become associated with the self would be a significant addition to our understanding of the self-concept and the extent to which a new voice may be perceived as self-relevant.

This question is clinically pertinent to people with voice impairments who may use another’s voice to communicate, i.e., within a text-to-speech device. Reports from users of such devices make it clear that the loss of their own self-voice can lead to a loss of self-identity (Nathanson, 2017). Therefore, exploring new ways of better incorporating a new self-voice into their concept of self is of direct importance. Indeed, it may help to elucidate what it is that makes a voice ‘mine’.

In the rest of the chapter, I present three experiments using the perceptual matching paradigm. Across these experiments I first assess whether the self-prioritisation effect is found for a new self-owned voice. Second, I begin to explore the factors that may facilitate it becoming self-relevant and therefore more closely associated with the self.

2.2 Experiment 1

In the first study, I asked whether the voice of another could become associated with the self through ownership, and prioritised in perception relative to voices associated with either a friend or a stranger. This experiment was preregistered via the Open Science Framework (https://osf.io/hg3w6/).

2.2.1 Hypothesis

Within the perceptual matching paradigm, I hypothesised that listeners would have quicker reaction times and greater recognition accuracy for their new self-voice compared to the voices associated with a friend or a stranger in match trials, thereby revealing a self-prioritisation effect.

2.2.2 Participants

All participants were recruited online via Prolific (www.prolific.ac) as native speakers of English with no hearing difficulties, and tested online using Gorilla (gorilla.sc, Anwyl-Irvine, Massonnié, Flitton, Kirkham, & Evershed, 2018). Participants also had to have an approval rate of over 90% to be eligible and were required to pass a headphone check (Woods, Siegel, Traer, & McDermott, 2017) to ensure they were wearing headphones and able to hear the stimuli. In Experiments 1-3, this headphone task was implemented in Gorilla and required participants to judge which of three presented tones was the quietest. In each set of three tones, one was presented 180° out of phase across the stereo channels. This makes it easy to distinguish the quietest tone when wearing headphones but difficult over loudspeakers due to phase-cancellation (Woods et al., 2017). Participants were asked to do this in six trials and only participants with >85% accuracy were eligible to continue.

None of the participants had taken part in any of the pilot studies associated with this project and, upon completion of the study, were paid for their participation. Ethical approval was obtained (SHaPS-2018-CM-029, UK) and informed consent was obtained from all participants prior to testing.

In Experiment 1, I recruited 35 participants (mean age = 27.25 years, SD = 6.41 years, age range = 18-40, 6 female, 29 male). The sample size was determined according to a previous study using the same paradigm by Schäfer et al. (2016). Their a priori power calculation suggested an effect size of $dz \geq .52$ could be achieved given $N = 31$, $\alpha = .05$, and a power of $1 - \beta = .80$. A true effect size of $dz = .51$ was subsequently found. After exclusion criteria were applied (see 2.2.5 Pre-processing), data from 31 participants were analysed in Experiment 1.

2.2.3 General Methods

The following section outlines the methods for all three studies, which use the same perceptual matching paradigm. Specifics to Experiment 1 are included thereafter.

2.2.3.1 Stimuli

Across all experiments, the stimuli included auditory exemplars of three male speakers and three female speakers each saying ‘hello’ in Southern Standard British English. The male-voiced exemplars were extracted from a larger pool of original recordings, which were obtained in sound-attenuated booths using desktop computers running Audacity (https://www.audacityteam.org/; RRID = SCR_007198) and either a Røde NT1A microphone (Røde Microphones, Sydney, Australia) or a Neumann TLM103 microphone (Neumann, Berlin, Germany). The female-voiced exemplars were recorded using a Rode NT-1A microphone via an RME Fireface UC audio interface. Recordings were digitised using Audacity (Audacity Team, 2019) as 16-bit, stereo WAV files, with sample rate 44100Hz. All stimuli were normalised for RMS amplitude using PRAAT (Boersma & Weenink, 2010).

The three male voices and three female voices were each selected from a wider pool of voices based on trait ratings obtained in pilot studies. In the first pilot study, a group of listeners (n = 15 participants, mean age = 30.33 years, SD = 6.63, age range = 18-

40 years, 8 female, 7 male) rated 30 auditory exemplars of ‘hello’ from six different male speakers for attractiveness (‘How attractive does this voice sound?’) and, separately, for trustworthiness (‘How trustworthy does this voice sound?’) on a 7-point Likert scale whereby 1 denoted ‘*not attractive/trustworthy at all’* and 7 denoted ‘*very attractive/trustworthy’*. I selected three male voices for use in these experiments based on their being well matched in attractiveness (mean ratings ± SDs = 4.6 ± 1.2, 4.5 ± 1.3, 4.8 ± 1.4) and trustworthiness (mean ratings ± SDs = 4.6 ± 1.2, 5.2 ± 1.1, 4.7 ± 1.2).

In the same pilot, participants were asked to rank the six different male voices in order of preference, 1 to 6, whereby 1 denoted ‘*The voice I would most like to have as my own*’ down to 6, which denoted ‘*The voice I would least like to have as my own*’. A Friedman Test was then run in the R environment using *stats* (R Core Team, 2013) to determine whether the six voices differed significantly in their mean ranking. The results indicated that the voices did differ significantly ($\chi2(5) = 23.82$, $p<.001$) so further Nemenyi post-hoc tests were run using *PMCMR* (Pohlert, 2014). The comparisons revealed that, of the original six voices, the three male voices used across these three experiments were well matched in their mean ranking with all p-values >.05.

In a second pilot study, six female voices were also trait rated and ranked for preference in the same way (n = 15 participants, mean age = 27.6 years, SD = 6.31, age range = 18-39 years, 6 female, 9 male). As with the male voice stimuli, the three female voices selected for use in Experiments 2 and 3 were chosen for being well matched for attractiveness (mean ratings ± SDs = 4.4 ± 1.0, 4.2 ± 1.4, 4.8 ± 1.4) and for trustworthiness (mean ratings ± SDs = 4.8 ± 1.2, 4.6 ± 1.3, 4.9 ± 1.0).

Finally, the same participants were asked to rank the six different female voices in order of preference. The ranking data was analysed with a Friedman Test in the R environment and the results indicated that the voices differed significantly ($\chi2(5) = 28.83$, $p<.001$). Further Nemenyi post-hoc tests were run and the comparisons revealed that, of the original six voices, the three female voices used in Experiments 2, 3, 4, and 5 were well matched in their mean ranking with all p-values >.05.

Experiment 1 used just the male voice stimuli, with two auditory tokens of "hello" per voice.

2.2.3.2 Procedure

The perceptual matching paradigm involved two stages: the familiarisation phase and the test phase, illustrated in Figure 1. The online tasks here described can be accessed and previewed online via Gorilla Open Materials at: https://app.gorilla.sc/openmaterials/45935.

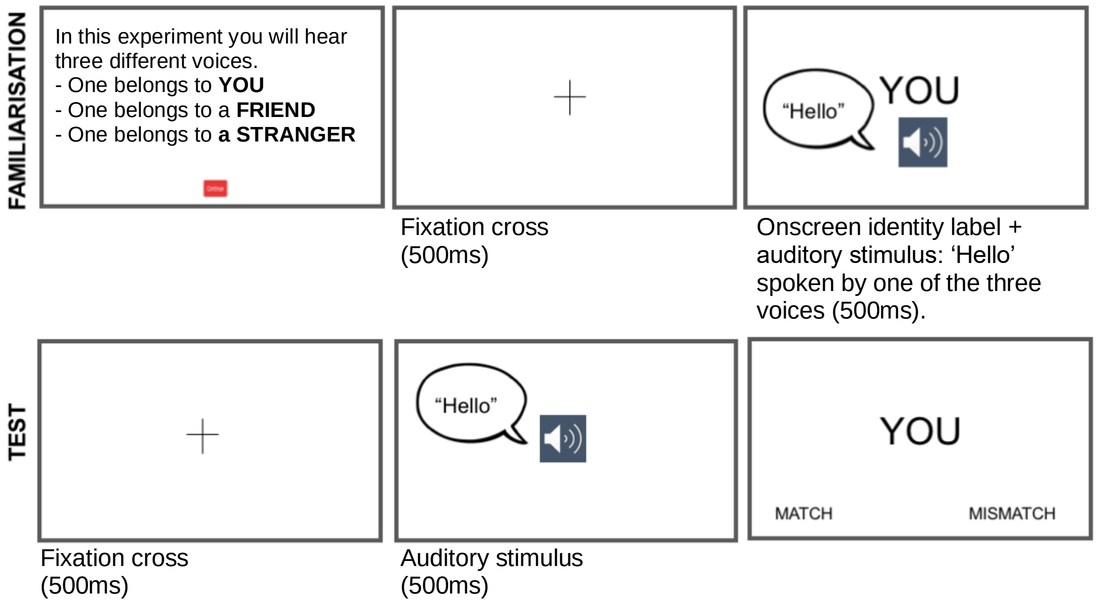

Figure 1. Experiment 1: The typical trial structure for the perceptual matching task (Sui et al., 2012) as used within Experiments 1-3, 6-7, and 10. After associating three new, unfamiliar voices with either 'self', 'friend', or 'other' identities, participants make speeded judgements about whether a voice they hear correctly matches the identity label presented on-screen or not. Faster reaction times and higher accuracy are indicative of greater perceptual prioritisation and self-bias.

1) Familiarisation Stage

In the familiarisation stage, participants were passively exposed to each of the three voices and their associated identity label. Following a 500ms fixation cross, the word 'YOU, 'FRIEND', or 'STRANGER' was displayed in the centre of the screen in black uppercase font on a white background and remained onscreen for 3000ms. 500ms after the label's onset, an auditory exemplar of 'hello' from the to-be-associated voice was played, lasting approximately 500ms. After the auditory stimulus finished, the label remained on the screen until the end of the 3000ms trial. The familiarisation phase consisted of 12 trials, with each label-voice pairing being presented four times, and each of the two auditory exemplars per voice played twice. Stimuli were presented in a random order. The familiarisation stage thus lasted approximately 1 minute, including on-screen instructions, and was immediately followed by the test phase.

2) Test Phase

The test phase started with 12 practice trials. Each trial started with a 500ms fixation cross in the centre of the screen, after which an auditory exemplar played for its total duration (approx. 500ms). Immediately after the auditory offset, the word ‘YOU’, ‘FRIEND’, or ‘STRANGER’ was displayed in the centre of the screen. Participants were asked to judge whether the identity label onscreen was a match or mismatch to the voice heard and to respond by keyboard press as quickly and as accurately as possible. Participants were instructed to press the left arrow for ‘MATCH’ and the right arrow for ‘MISMATCH’. This left-right ordering of match and mismatch remained constant on all trials, for all participants. In the practice trials only, text appeared on-screen to remind participants that a left button press meant a match, and a right button press meant mismatch. In all test trials, however, no visual reminder was provided. Feedback was given onscreen for 500ms immediately following every response: a green tick for correct, a red cross for incorrect, and text feedback of ‘TOO SLOW’ for responses occurring after 1500ms. The next trial began after the 500ms feedback period was complete.

Following the practice trials, participants performed three blocks of 72 trials (216 trials in total) and were informed of their overall accuracy at the end of each block. The order of match vs. mismatch trials was randomised for all participants. The experiment lasted approximately 25 minutes in total and participants were debriefed at the end.

2.2.3.3 Design

Experiment 1 consisted of two within-subject factors of ‘voice identity’ (*self* vs. *friend* vs. *other* voice) and ‘trial type’ (*match* vs. *mismatch*).

2.2.3.4 Pre-processing

In each experiment (1-3) I analysed three measures: reaction times, sensitivity, and accuracy. The first two measures were included to align with previous reports of the self-prioritisation effect by Sui et al. (2012) and Schäfer et al. (2016). I also included trialwise accuracy to make it possible to run a complementary analysis to the RT mixed model that included the same random effects structure.

Reaction times: RTs were measured as the delay between the onset of the visual label and the participant's response and categorised by the *voice identity* (self-voice, friend-voice, other-voice) per trial, and not by the category of the presented label ("YOU", "FRIEND", "STRANGER"). All erroneous responses, as well as responses shorter than 200ms or longer than 1500ms were excluded from the RT analysis in line with Sui et al.'s (2012) approach.

Sensitivity: unbiased *d'* scores were calculated by combining performance scores from match and mismatch trials at each level of the *voice identity* factor (self, friend, other). Specifically, YES/NO responses were recoded into hits and false alarms, following the log-linear approach to adjust for cases involving 100% hits or 0% false alarms (Stanislaw & Todorov, 1999).

Accuracy: Trialwise accuracy scores were recorded for each participant. Whole datasets for participants showing overall performance accuracy at chance (≤50% + 95% CI) were excluded as this indicated random responses.

2.2.3.5 Exclusion

In Experiment 1, 2.48% of the trials were eliminated on the RT criteria above as well as whole datasets for two participants with more than 50% of responses shorter than 200ms. A further two datasets were eliminated on the accuracy criterion. None of these participants was replaced, therefore data from a total of 31 participants were analysed in Experiment 1.

2.2.3.6 Analysis

Across Experiments 1-3, reaction times and sensitivity were assessed with linear mixed models (LMM) using *lme4* (Bates, Maechler, Bolker & Walker, 2014) in the *R* environment (R Core Team, 2013). Accuracy was assessed using binomial generalised linear mixed models (GLMM). In every analysis (with the exception of sensitivity in Experiment 1), I ran a model that included an interaction between *voice identity* and the manipulated variable, two fixed effects (*voice identity*, and the manipulated variable: *trial type* in Experiment 1, *gender-matching* in Experiment 2, and *choice* in Experiment 3), and a random intercept of 'participant'.

Statistical significance of the effects was established via likelihood ratio tests by dropping effects of interest from the appropriate model. For example, to establish whether the interaction was significant, I dropped this interaction from the model including both fixed effects. To test for the significance of either of the main effects, I dropped the relevant effects from a model that only included the two main effects. For all analyses reported here, post-hoc comparisons were conducted in *emmeans* (Lenth, 2016) and were adjusted for the multiple comparisons via Bonferroni correction.

For GLMMs, odds ratios and associated confidence intervals are provided as an estimate of the size of the relevant effects. An odds ratio of 1 indicates that no effect is present, while odds ratios that deviate from 1 indicate an effect is present. The larger the deviation, the bigger the effect. If the confidence intervals do not cross 1, the relevant effect is significant. For LMMs, the models' estimates and associated confidence intervals for each effect are reported. The further away from 0 the estimates are, the bigger the effect. If the confidence intervals do not cross 0, the relevant effect is significant.

2.2.4 Results & Discussion

Descriptive statistics for reaction times, accuracy, and sensitivity scores (d') for the different conditions and voice identities are given in Table 1.

Table 1*. Mean RTs, accuracy, and sensitivity (d') for Experiment 1.*

Condition	**Voice Identity**	**Mean RT (ms)**	**Accuracy**	***d' ****
Match trials	Self	584 (211)	0.91 (0.29)	2.64 (1.11)
	Friend	635 (220)	0.88 (0.33)	2.31 (0.97)
	Other	645 (227)	0.87 (0.34)	2.29 (1.06)
Mismatch trials	Self	714 (242)	0.83 (0.38)	
	Friend	722 (217)	0.81 (0.39)	
	Other	712 (216)	0.81 (0.40)	

Note. RT = reaction time; Accuracy = proportion correct. Standard deviations appear within parentheses. *Performance scores from match and mismatch trials are combined to provide d' scores.

2.2.4.1 Reaction times

The reaction time data are plotted in Figure 2. I ran an LMM on RTs to assess whether the self-voice had been perceptually prioritised relative to the others depending on the *trial type* (match trials vs mismatch trials). For the post-hoc tests, alpha was set as .008 after Bonferroni correction for six comparisons.

E.g. lmer(reaction time ~ voice identity * trial type
+ 1 | participant, REML = "FALSE"

Full model outputs are reproduced in Supplemental Table 1 (see Appendix). There was a significant interaction between *voice identity* and *trial type* ($\chi 2(2)= 21.32$, $p<.001$). Post-hoc tests confirmed that reaction times for the self-voice were significantly quicker than those for the friend-voice ($p<.001$) and the other-voice ($p<.001$). However, this was only the case within matched trials, which is in line with previous studies (Sui et al., 2012). There was no significant difference between reaction times for the friend-voice and other-voice in either match, or mismatch trials ($ps > .008$). These results show therefore a significant self-prioritisation effect for the perception of voices, thus supporting the hypothesis.

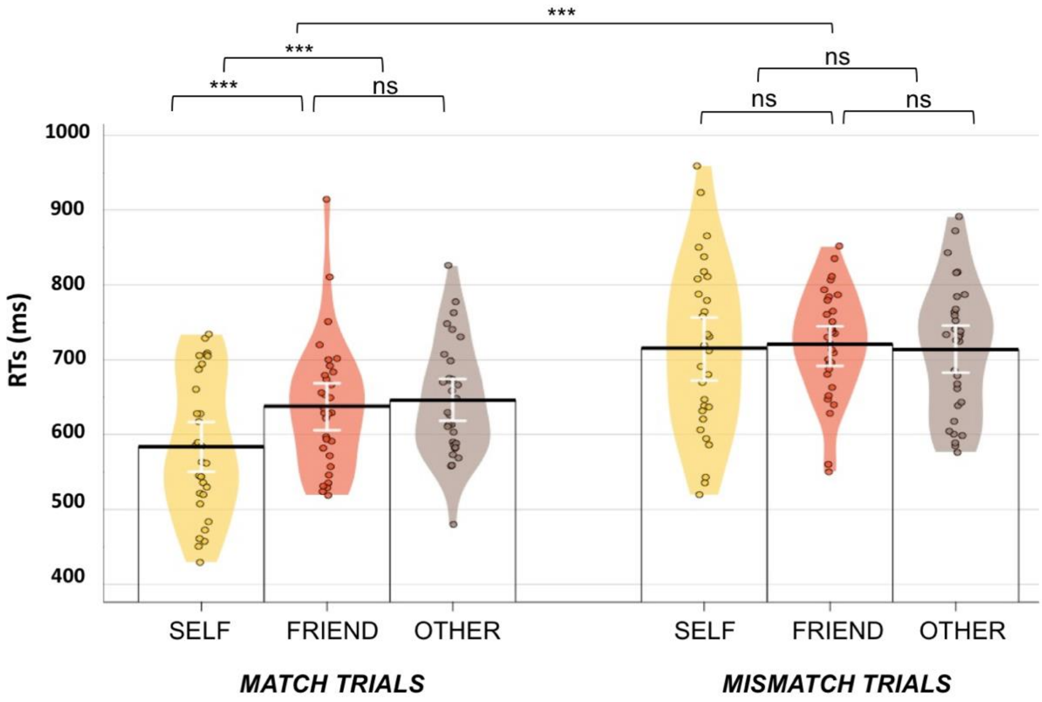

Figure 2. Experiment 1: Mean reaction times (RT) as a function of the voice identity (self, friend, or other) and the trial type (match vs mismatch). The error bars indicate the SEs of the means. Coloured segments show smooth density curves for the full data distribution, while individual dots indicate mean RTs per participant. The top bar indicates the significance of the interaction, horizontal bars show post-hoc comparisons. Asterisks denote significance within RTs as determined by likelihood ratio tests. *** p<.001

2.2.4.2 Sensitivity

I also ran an LMM on *d'* scores to assess performance in the perceptual matching task using the model below. For the post-hoc tests, alpha was set as .017 after Bonferroni correction for three post-hoc comparisons.

E.g., lmer (sensitivity ~ voice identity + 1 | participant, REML = "FALSE")

Full model outputs are reproduced in Supplemental Table 2. There was again a significant main effect of *voice identity* ($\chi2(2)=6.24$, $p=.044$). However, post-hoc comparisons showed that although sensitivity for the self-voice was numerically greater than that for both the friend-voice and other-voice, the difference did not reach significance (ps >.017) after Bonferroni correction.

2.2.4.3. Accuracy

Finally, I ran a GLMM on trialwise accuracy. For the post-hoc tests, alpha was set as .017 after Bonferroni correction for three post-hoc comparisons.

$$\text{E.g. glmer(accuracy} \sim \text{voice identity * trial type} \\ + \ 1 \mid \text{participant, family = "binomial")}$$

Full model outputs are reproduced in Supplemental Table 3. The interaction of *voice identity* and *trial type* was non-significant ($\chi2$ (2)= 2.29, p=.318). There was, however, a significant main effect of *voice identity* ($\chi2(2)$=11.84, p=.003), showing that performance accuracy for the self-voice was higher than for both the friend-voice (p=.013) and the other-voice (p<.001), while the friend-voice and other-voice did not significantly differ. There was also a main effect of *trial type* ($\chi2(1)$=70.63, p<.001) showing that accuracy was significantly higher in match trials than in mismatch trials (p<.001); a result which falls in line with previous studies.

In Experiment 1, I asked whether an unfamiliar voice that, inherently, belongs to another could be made self-relevant through ownership and accepted as part of the self-concept. The results reveal a clear self-prioritisation effect, with quicker reaction times, higher accuracy, and increased perceptual sensitivity to the newly self-owned voice relative to the other voice identities. This shows that the new self-voice has been attributed self-bias and processed as a self-relevant stimulus. This therefore supports the hypothesis and provides a pattern of results that mirrors those reported by Sui et al. (2012), Schafer et al. (2016), and Payne et al. (2017). This result demonstrates for the first time that what is ‘self’ can quickly be extended to incorporate an unfamiliar voice that is self-owned but that, inherently, signals a different identity. The fact that such a stimulus can be processed as self-relevant further shows that a new voice does not need to be self-generated in order to be deemed “self” and perceptually prioritised.

2.3 Experiment 2

2.3.1 Introduction

Having demonstrated that there is a self-prioritisation effect in the perception of voices, Experiment 2 asked whether this effect is increased if the new self-voice is matched to the gender-identity of the listener. This experiment was preregistered via the Open Science Framework (https://osf.io/r6g2m/) and can be previewed via Gorilla Open Materials at: https://app.gorilla.sc/openmaterials/46084.

The self-voice plays a key role in representing the self-identity to others. It is both a conduit through which thoughts and feelings can be expressed and, importantly here, a signal that conveys a unique identity – an embodied self. Indeed, the sound of one's voice is partly determined by - and wholly representative of – a person's body-size, sex, health, age as well as social information such as regional origin and economic status (Sidtis & Kreiman, 2011). Each voice is unique to the individual. Thus, a new self-voice than is more closely matched to that individual – such that it may be a more accurate means of self-representation – may be deemed more self-relevant.

How self-relevant the voice is may influence prioritisation, such that is it relatively more prioritised if more relevant. Indeed, Golubickis, Falbén, Ho, Sui, Cunningham, & Macrae (2019), demonstrated that self-prioritisation is sensitive to how relevant stimuli are to the self-identity. Specifically, they used an adapted perceptual matching paradigm and asked participants to form new associations between the classic geometric shapes (circle, square, triangle) and, newly, three identity-relevant groups to which the participant felt they belonged. Importantly, one group was of high identity relevance (i.e., musician), another of medium identity relevance (i.e., vegan) and one of low identity relevance (i.e., athlete). At test, the shape that had been associated with the group of the highest identity-relevance was prioritised in comparison to the other shapes. This shows that, despite all three shapes becoming self-relevant (in that they were associated to three in-groups), the degree of prioritisation was modulated according to the degree of self-relevance.

More widely, previous studies have demonstrated that people are more biased towards voices they perceive to be similar to their own (Peng, Wang, Meng, Liu, Hu, 2019) and, indeed, towards their own voice (Hughes and Harrison, 2013). Both the self-voice and voices perceived as similar to self are judged to be more attractive than voices of dissimilar others. More generally still, people prefer others whom they perceive or imagine to be similar to themselves (Gerson, Bekkering & Hunnius, 2017;

Hamlin, Mahajan, Liberman, & Wynn, 2013; Mahajan & Wynn, 2012) because it activates the positive and implicit associations they already hold about themselves (Jones, Pelham, Carvallo, & Mirenberg, 2004). The perceived similarity reduces the social distance between them and, relatedly, reduces the degree to which the self and the other are differentially processed (Benoit, Gilbert, Volle, & Burgess, 2010). However, whether this bias in judgement also generalises to prioritisation in perception is unclear.

Overall though, it is reasonable to suggest that an other-generated voice that is more similar to the true self-voice of the listener - in terms of gender-related acoustics - may be perceived as more closely related to the self than a voice which represents a less similar other. Further, an externally-generated vocal signal that fits better within the parameters of our own vocal apparatus and better aligns with our identity may be deemed more personally relevant than a voice which embodies and represents a different gender identity.

2.3.2 Hypothesis

I therefore hypothesised that the self-prioritisation effect for the self-voice would increase when it was matched to the gender-identity of the participant. Such an increase should result in quicker reaction times and higher accuracy in response to the self-voice, relative to the friend- or other-voice, when the self-voice is matched to the gender identity of the participant as compared to when the self-voice is not matched to the gender identity of the participant.

2.3.3 Participants

96 participants took part (48 females: mean age = 29.1 years, SD = 6.28 years, age range = 18-40; 48 males: mean age = 27.4 years, SD = 6.58 years, age range = 18-40). All participants were cisgender such that their gender-identity aligned with their sex assigned at birth. The data from participants (N=2) showing overall performance accuracy at chance (≤50% + 95% CI) were excluded and replacement participants sampled such that data from 96 participants were included in the reported analyses.

2.3.4 Methods

Unless stated below, the methods for Experiment 2 replicated methods from Experiment 1.

2.3.4.1 Randomisation

Male and female participants were randomly allocated to either a ‘gender-matched’ group (N=48: 24M, 24F) or a ‘non-gender-matched’ group (N=48: 24M, 24F), whereby the gender of the participant would match or not match the gender of the voices.

2.3.4.2 Stimuli

In addition to the male voice stimuli used in Experiment 1, I included female voice stimuli from three SSBE-accented speakers.

2.3.4.3 Design

The study consisted of a within-subjects factor of ‘voice identity’ (*self* vs. *friend* vs. *other*) and a between-subjects factor of ‘gender-matching’ (*gender-matched* vs *gender-mismatched*). Further, a within-subjects factor of trial type was present (*match* vs *mismatch* trials).

2.3.4.4 Pre-processing

Based on the exclusion criteria reported earlier, fewer than 1% of the trials were eliminated.

2.3.4.5 Analysis

In Experiment 1 I found that the self-prioritisation effect emerged on match trials only (i.e., trials in which the presented voice and identity label are correctly ‘matched’ according to their learnt association). Therefore, following the logic of Sui et al. (2012) and Schafer et al. (2016), the analyses of reaction times and raw accuracy are henceforth reported separately for match and mismatch trials.

2.3.5 Results and Discussion

Descriptive statistics for reaction times, accuracy, and sensitivity scores (*d’*) for the different conditions and voice identities are given in Table 2.

***Table 2**. Mean RTs, accuracy and sensitivity (d') for Experiment 2 (match trials).*

Condition	**Voice Identity**	**Mean RT (ms)**	**Accuracy**	***d'* ***
Gender-matched	Self	541 (203)	0.92 (0.27)	2.89 (0.98)
	Friend	572 (213)	0.92 (0.27)	2.90 (1.01)
	Other	592 (221)	0.85 (0.35)	2.49 (0.97)
Gender-mismatched	Self	536 (205)	0.94 (0.24)	2.95 (1.04)
	Friend	570 (222)	0.91 (0.29)	2.87 (1.16)
	Other	589 (224)	0.88 (0.33)	2.68 (1.21)

Note. RT = reaction time; Accuracy - proportion correct. Standard deviations appear within parentheses. *Performance scores from match and mismatch trials are combined to provide *d'* scores.

2.3.5.1 Reaction times

The reaction time data are plotted in Figure 3. As in the analysis in Experiment 1, I ran an LMM on RTs to assess whether the self-voice had been perceptually prioritised. First, I analysed match trials only. For the post-hoc tests, alpha was set as .017 after Bonferroni correction for three post-hoc comparisons.

E.g. lmer(reaction time ~ voice identity * gender-matching
+ 1 | participant, REML = "FALSE")

Full model outputs are reproduced in Supplemental Table 4. The interaction of *voice identity* and *gender-matching* was non-significant ($\chi2(2)= .27$, $p = .875$). There was, however, again a main effect of *voice identity* ($\chi2(2)=106.78$, $p<.001$): in both groups, reaction times for the self-voice were significantly quicker than those for either the friend-voice ($p<.001$) or the other-voice ($p<.001$), and reaction times to the friend-voice were also significantly quicker than those for the other-voice ($p<.001$). Lastly, I found no main effect of *gender-matching* on reactions times ($\chi2(1) < .01$, $p=.934$).

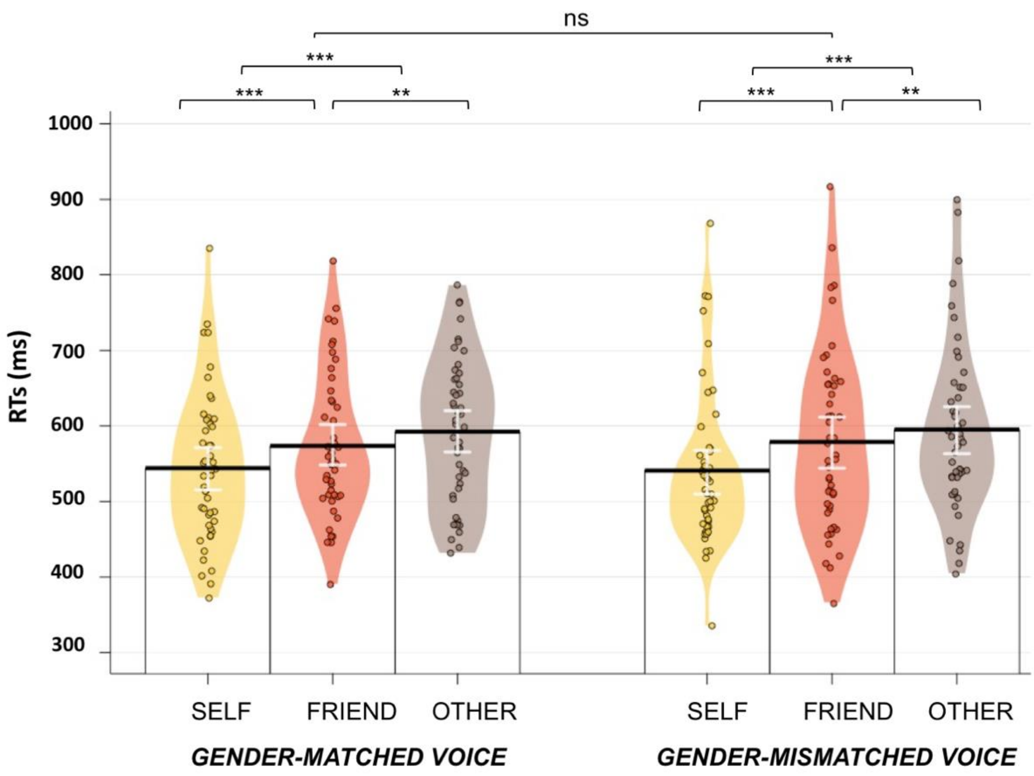

Figure 3. Experiment 2**:** Mean reaction times (RT) as a function of the voice identity (self, friend, or other) and gender-matching (matched vs mismatch, i.e., whether the voice matched the gender-identity of the participant, or not). **Graph models match trials only.** The error bars indicate the SEs of the means. Coloured segments show smooth density curves for the full data distribution, while individual dots indicate mean RTs per participant. The top bar indicates the significance of the interaction, horizontal bars show post-hoc comparisons. Asterisks denote significance within RTs as determined by likelihood ratio tests. *** p<.001

These RT data therefore reveal a significant self-prioritisation effect in the perception of voices, but show that prioritisation was not modulated according to whether the self-voice was matched to the gender-identity of the participant.

Identical analyses of RT data were then run on the mismatch trials, but the interaction ($\chi2(2)$= 1.03, p=.596), the effect of *voice identity*, ($\chi2(2)$= .07, p=.968), and the effect of *gender matching* ($\chi2(2)$= .69, p=.708) were all non-significant in mismatch trials.

2.3.5.2 Sensitivity

I also ran an LMM on *d'* scores. For the post-hoc tests, alpha was set as .017 after correction for three post-hoc comparisons.

E.g. lmer(sensitivity ~ voice identity * gender-matching
+ 1 | participant, REML = "FALSE")

Full model outputs are reproduced in Supplemental Table 5. The interaction of *voice identity* and *gender-matching* was non-significant ($\chi2(2) = 1.57$, $p = .455$). There was again a main effect of *voice identity* ($\chi2(2) = 16.62$, $p<.001$), with post-hoc comparisons revealing that, across groups, sensitivity for the self-voice was significantly higher than for the other-voice ($p<.001$) but did not differ significantly from the friend-voice. Further, sensitivity for the friend-voice were significantly higher than for other-voice ($p<.001$). There was again no effect of *gender-matching* on reaction times ($\chi2(1)= .16$, $p= .687$).

2.3.5.3 Accuracy

Finally, I ran a GLMM on trialwise accuracy. Here I analysed the match trials. For the post-hoc tests, alpha was set as .008 after Bonferroni correction for six comparisons.

E.g. glmer(accuracy ~ voice identity * gender-matching
+ 1 | participant, family = "binomial")

Full model outputs are reproduced in Supplemental Table 6. There was a significant interaction between *voice identity* and *gender-matching* on accuracy ($\chi2(2)=7.09$, $p= .029$). Post-hoc comparisons revealed that accuracy for the self-voice was significantly higher than accuracy for the other-voice ($p<.001$) in both gender-matched and non-gender-matched groups. However, accuracy for the self-voice only differed from the friend-voice in the non-gender-matched group ($p=.001$). Accuracy for the self-voice did not differ significantly when it was gender-matched in comparison to when it was not.

Identical analyses within the mismatch trials revealed that neither the interaction effect ($\chi2(2)$= 1.03, p=.596), the fixed effect of *voice identity* ($\chi2(2)$=.065, p=.968), nor the fixed effect of *gender-matching* ($\chi2(1)$=.069, p= .793) were significant.

In Experiment 2, I asked whether the self-prioritisation effect could be modulated by the self-voice being gender-matched to the identity of the listener. The results show that *gender-matching* did not significantly influence prioritisation here, with equivalent reaction times and perceptual sensitivity to the voices when they were matched as to when they were mismatched. Although there was a significant interaction of *voice identity* and *gender-matching* for accuracy scores, this did not manifest in interpretable differences in terms of how the self-voice related to the friend-voice and the other-voice across the two groups. Overall, these results suggest that the perceived physiological similarity between the novel voices and the participant's own voice had no effect on the self-prioritisation effect. This is unexpected (see the General Discussion, section 2.5, for suggested explanations). Instead, prioritisation was similar for both groups, in that perceptual processing was biased towards the new self-voice relative to the friend- or other-voice, thereby supporting the results from Experiment 1.

2.4. Experiment 3

2.4.1 Introduction

Experiment 3 further investigated which factors may influence the acceptability of an externally-generated voice being integrated into the self-concept. Specifically, I here asked whether the action of personally choosing an externally-generated voice for the self increases the level of bias towards it. This experiment was preregistered via the Open Science Framework (https://osf.io/wz2am/) and can be previewed via Gorilla Open Materials at https://gorilla.sc/openmaterials/46086.

In Experiments 1 and 2, the new self-voice became associated with the self via ownership. That is, participants were told the voice belonged to them and the results demonstrated that the voice was subsequently prioritised in perception, having become self-relevant as a self-owned stimulus. It is possible, however, that having personal choice over what is 'self' may increase bias towards it. Indeed, previous studies have shown that having a choice is, itself, inherently rewarding (Leotti &

Delgado, 2015) and a voice that has been chosen to represent the self may be perceived as more self-relevant. As reviewed in Chapter 1, Huang et al., (2009) demonstrated that, of two identical items, people will prefer the item that they have chosen and value it more highly.

The effect of choice and, relatedly, of ownership, on self-bias has been investigated in the memory domain by Cloutier and Macrae (2008). They demonstrate that participants who are given a choice over which items to own can better recall those items relative to participants who had no choice and were instead afforded ownership by someone else. The act of selection, of making a personal choice over external stimuli, is associated with enhanced memory for the later outcomes of those choices (Cloutier & Macrae, 2008). According to the Cunningham et al., (2008) this may be underpinned by stronger encoding for self-relevant information. Furthermore, it is possible that being given a choice may enhance the participant's sense of having agency in extending the self to include that new voice.

2.4.2 Hypothesis

Experiment 3 therefore predicted that the self-prioritisation effect in voices would be increased if the new, self-voice was first chosen by the participant, relative to when it was randomly allocated. Such an increase in prioritisation should result in quicker reaction times and higher accuracy in response to the self-voice – in comparison to the voices associated with the friend or other – when the new self-voice was chosen, relative to when it was randomly allocated.

2.4.3 Participants

96 new participants were recruited, forming a 'choice' group (48 females: mean age = 29.2 years, SD = 6.27 years, age range = 18-40; 48 males: mean age = 26.5 years, SD = 6.71 years, age range = 18-40). As in previous analysis, datasets for participants showing overall performance accuracy at chance (≤50% + 95% CI) were excluded (N=3), as this indicated random responses. These participants were replaced to reach the full sample size of 96 participants. Data from these 96 participants were then combined with data from the 96 participants from Experiment 2 who had been randomly allocated voices, thus forming the 'no choice' group (48 females: mean age

= 29.1 years, SD = 6.28 years, age range = 18-40; 48 males: mean age = 27.4 years, SD = 6.58 years, age range = 18-40). Thus, data from 192 participants were included in this study.

2.4.4 Methods

Unless stated below, the methods for Experiment 3 replicated methods from Experiments 1 and 2.

2.4.4.1 Stimuli

Experiment 3 used both male and female voice stimuli, as in Experiment 2.

2.4.4.2 Procedure

For the ‘choice’ group, Experiment 3 began with a voice selection task in which the participants were presented with two tokens of ‘hello’ from each of the three male or female voices, labelled onscreen as ‘Voice A’, ‘Voice B’, and ‘Voice C’. They were then asked to choose which voice they wanted to represent them for the rest of the study. The voices for ‘friend’ and ‘other’ were allocated randomly, and evenly, thereafter. It is important to note, however, that these two voices were effectively rejected as a self-voice. Second, participants were asked to rate the strength of their preference for their chosen voice on a 5-point scale where 1 denoted, ‘*Not strong at all: I was indifferent/picked randomly*’ and 5 denoted, ‘*Very strong: I preferred my chosen voice much more than the others*.’ Thereafter, the procedure replicated that of Experiments 1 and 2.

2.4.4.3 Design

The study consisted of a factorial design with one within-subjects factor, ‘voice identity’ (*self* vs. *friend* vs. *other*) and one between-subjects factor, ‘choice’ (*choice* vs. *no choice*). Further, a within-subjects factor of ‘trial type’ was present (*match* vs *mismatch* trials).

2.4.4.4 Pre-processing

In the self-voice selection task, participants were presented with a choice of 3 male voices, or 3 female voices. For participants choosing from male voices, there was a relatively balanced selection across the three identities. Voice 1 was chosen by 37.5% of participants (8F, 10M), Voice 2 was also chosen by 37.5% of participants (8F, 10M), and Voice 3 was chosen by 25% of participants (8F, 4M). The strength of preference task showed that 87.5% of participants stated they had a preference strength of at least three (out of five) for the voice they chose, with only 2.08% stating that they were indifferent to their self-voice.

Uptake of the female voices as the new self-voice was less balanced; Voice 1 was chosen by 62.5% of participants (16F, 15M), Voice 2 was chosen by 12.5% (4F, 2M), and Voice 3 was chosen by 25% of participants (4F, 7M). The strength of preference task showed that 95.8% of participants stated they had a preference strength of at least three (out of five) for the voice they chose. Further, no participants stated that they were indifferent to their self-voice.

Since I posited that encoding of a stimulus to the self is strengthened through making a personal choice – rather than by the strength of preference for that stimulus – data from all participants were analysed regardless of their stated preference strength.

1.18% of the trials were eliminated from the analysis of RTs based on our exclusion criteria.

2.4.4.5 Analysis

For the analyses of reaction times and raw accuracy, results within match and mismatch trials are reported separately, following the logic of Sui et al. (2012) and Schafer et al. (2016). Match trials are those in which the presented voice and identity label are correctly 'matched' according to their learnt association.

2.4.5 Results & Discussion

Descriptive statistics for reaction times, accuracy, and sensitivity scores (*d'*) for the different conditions and voice identities are given in Table 3.

Table 3. *Mean RTs, accuracy, and sensitivity (d') in Experiment 3 (match trials).*

Condition	**Voice Identity**	**Mean RT (ms)**	**Accuracy**	***d' **
Choice	Self	542 (204)	0.94 (0.25)	3.10 (1.06)
	Friend	610 (231)	0.90 (0.30)	2.83 (1.15)
	Other	606 (234)	0.87 (0.33)	2.66 (1.19)
No Choice	Self	539 (204)	0.93 (0.27)	2.92 (1.01)
	Friend	571 (218)	0.92 (0.28)	2.89 (1.09)
	Other	590 (222)	086 (0.34)	2.59 (1.10)

Note. RT = reaction time; Accuracy - proportion correct. Standard deviations appear within parenthesis. *Performance scores from match and mismatch trials are combined to provide *d'* scores.

2.4.5.1 Reaction times

The reaction time data are plotted in Figure 4 for the match trials and in Figure 5 for the mismatch trials. I ran an LMM on RT data to assess whether the self-voice had been perceptually prioritised relative to the others depending on choice (i.e. whether the self-voice was chosen or not). First, I analysed match trials only. For the post-hoc tests, alpha was set as .008 after Bonferroni correction for six comparisons.

E.g. lmer(reaction time ~ voice identity * choice
+ 1 | participant, REML = "FALSE")

Full model outputs are reproduced in Supplemental Table 7. There was an interaction between *voice identity* and *choice* ($\chi2(2)=27.56$, ($p<.001$). I hypothesised that an interaction would arise because of a difference in RTs to the self-voice when it is chosen in comparison to when it is not. However, instead, the interaction arose from a difference in RTs to the friend-voice when the self-voice was chosen in comparison to when it was not. Specifically, when the self-choice was chosen, the RTs for the friend-voice increased such that they became equivalent to those of the other's voice ($p=.577$). Thus, the small bias that is typically shown for the friend-voice relative to the other-voice was here diminished. This may have been because the friend-voice had been actively rejected as a self-voice in the voice selection task. This left the self-voice relatively more prioritised than either of the other two voices, which is not the case when the self-voice is randomly allocated. When the self-voice was randomly allocated and none of the voices was actively rejected as a self-voice, reaction times to friend-voice (p<.001) and other-voice (p<.001) were significantly different. That is, the perceptual distance between the self-voice and the friend voice was reduced.

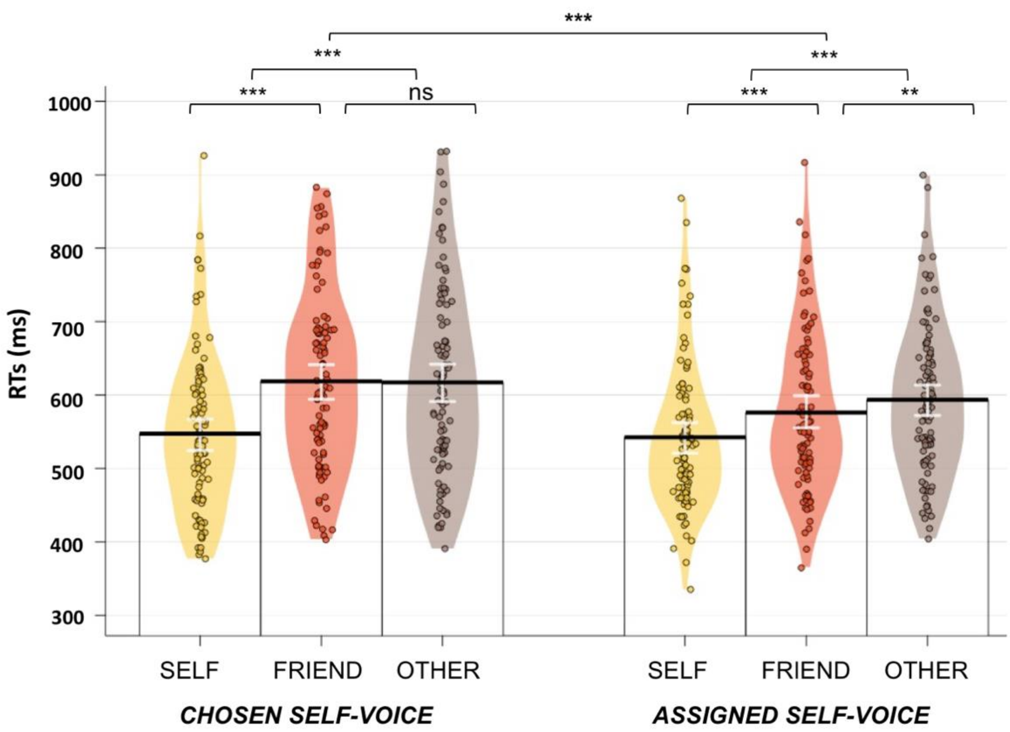

Figure 4. Experiment 3**:** Mean reaction times (RT) as a function of the voice identity (self, friend, or other) and choice (choice vs no choice, i.e., whether the self-voice was personally chosen or randomly allocated). **Graph models matched trials only.** The error bars indicate the SEs of the means. Coloured segments show smooth density curves for the full data distribution, while individual dots indicate mean RTs per participant. The top bar indicates the significance of the interaction, horizontal bars show post-hoc comparisons with asterisks denoting significance within RTs as determined by likelihood ratio tests. *** p<.001

Identical analyses were run on RTs in mismatch trials (Fig 5). For the post-hoc tests, alpha was set as .008 after Bonferroni correction for six comparisons. These analyses also showed a significant interaction between *voice identity* and *choice* ($\chi 2(2)=10.45$, $p=.005$). Reaction times were significantly quicker for the self-voice than for either the friend-voice (p<.001) or the other-voice (p=.001), but this self-prioritisation effect was only present when the self-voice was chosen. When the self-voice was randomly allocated, reaction times did not significantly differ between the self-, friend- and other-voice. This result supports the notion that the degree to which a self-voice is prioritised relative to others is influenced by whether it has been chosen or rejected as a self-owned item.

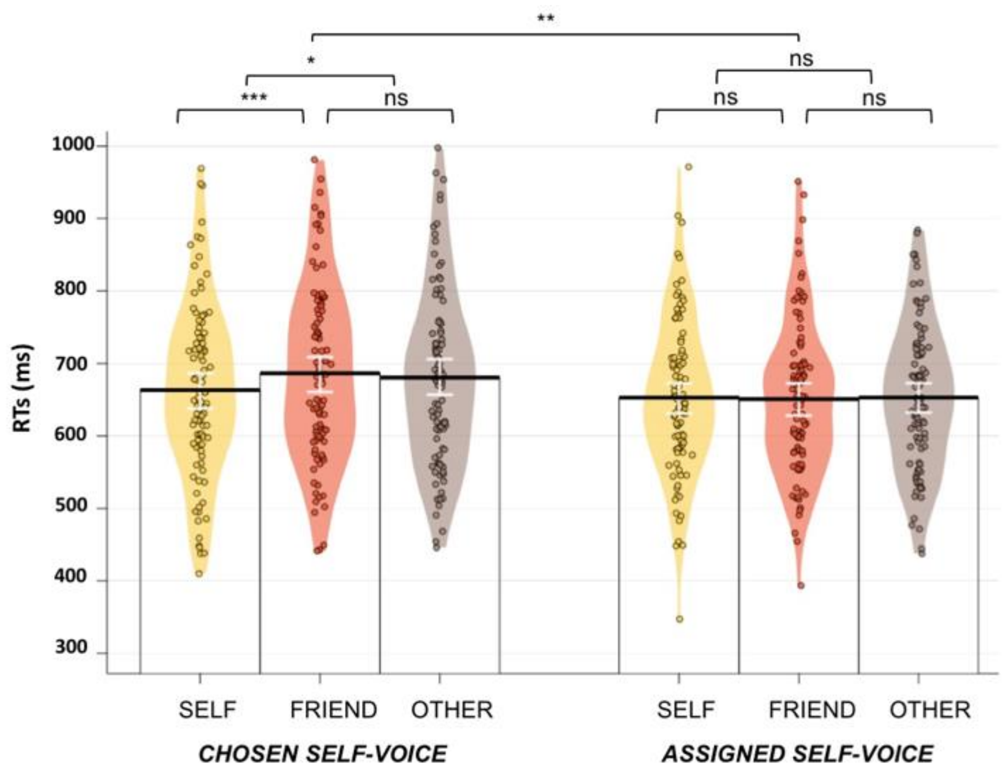

Figure 5. Experiment 3**:** Mean reaction times (RT) as a function of the voice identity (self, friend, or other) and choice (choice vs no choice, i.e., whether the self-voice was personally chosen or randomly allocated) **in mismatch trials only**. The error bars indicate the SEs of the means. Coloured segments show smooth density curves for the full data distribution, while individual dots indicate mean RTs per participant. The top bar indicates the significance of the interaction, horizontal bars show post-hoc comparisons. Asterisks denote significance within RTs as determined by likelihood ratio tests. *** p<.001

2.4.5.2 Sensitivity

I then ran an LMM on *d'* scores. For the post-hoc tests, alpha was set as .017 after Bonferroni correction for three post-hoc comparisons.

E.g. lmer(sensitivity ~ voice identity * choice

+ 1 | participant, REML = "FALSE")

Full model outputs are reproduced in Supplemental Table 8. The interaction of *voice identity* and *choice* was non-significant ($\chi2(2) = 3.55$, $p = .169$). There was, however, a main effect of *voice identity* on *d'* scores ($\chi2(2)= 34.55$, $p<.001$), revealing that sensitivity for the self-voice was significantly higher than for the other-voice ($p<.001$) but did not differ significantly from the friend-voice ($p=.020$). Further, participants' sensitivity also differed significantly between the friend-voice and other-voice ($p<.001$). Lastly, I tested for the significance of *choice* as a fixed effect, however this was non-significant ($\chi2(1) = .26$, $p= .612$).

2.4.5.3 Accuracy

Finally, I ran GLMMs on trialwise accuracy, first in the match trials. For the post-hoc tests, alpha was set as .008 after Bonferroni correction for six comparisons.

E.g. glmer(accuracy ~ voice identity * choice
+ 1 | participant, family = "binomial")

Full model outputs are reproduced in Supplemental Table 9. There was an interaction between *choice* and *voice identity* ($\chi2(2)=6.56$, $p=.038$). Pairwise comparisons revealed that when the self-voice was chosen, accuracy for the self-voice became significantly greater than accuracy for either the friend-voice ($p<.001$) or the other-voice ($p<.001$) showing a clear self-prioritisation effect. However, when the self-voice was randomly allocated, accuracy for the self-voice did not differ significantly from the friend-voice ($p=.020$) though it did differ from the other-voice ($p<.001$). Thus, the interaction effect arose from a relative deprioritisation of the friend-voice when the self-voice was chosen compared to when it was not. This pattern of results mirrors the interaction effect found in the RTs: here too, performance for the friend-voice is reduced in the choice condition, leaving the chosen self-voice relatively more prioritised than either of the other two voices. This is not the case when the self-voice was randomly allocated. Lastly, across both groups, accuracy for the friend-voice was significantly higher than accuracy for the other-voice (ps<.001).

Identical analyses were also run on accuracy in mismatch trials. For the post-hoc tests, alpha was set as .008 after Bonferroni correction for six comparisons. There was a significant interaction effect here too ($\chi2(2)=7.98$, $p=.018$). The interaction effect arose from significantly higher accuracy for the self-voice than for the other-voice ($p=.005$), only when the self-voice was chosen. When the self-voice was randomly allocated, performance did not significantly differ for the self, friend or other-voice.

2.4.6 Exploratory Analyses

As a means to further explore the implication of choice on self-prioritisation, I indexed the degree of bias for the self-voice (following Sui & Humphreys, 2017) in both the choice and no choice conditions. A larger score indicates a larger bias, whereas smaller scores indicate a relatively smaller bias. For each participant, a self-bias score can be calculated by taking the difference in RTs for the self-voice versus the friend-voice, dividing it by the sum of RTs across the two conditions, and finally multiplying the result by 100 to achieve a percentage score. A friend-bias score can also be calculated by taking the difference in RT performance for the friend-voice relative to the other-voice, dividing it by the sum of the RTs across the conditions and multiplying it by 100. Bias scores are plotted in Figure 6.

The bias scores show that there is a larger bias – and greater prioritisation – for 'self' relative to 'friend' across both conditions. Importantly though, when the self-voice had been chosen – and, ipso facto, the other voices had been rejected – self-prioritisation increased, and the prioritisation of friend decreased. This supports the earlier findings; that there is a greater prominence of self, relative to others, when the stimulus is personally selected. In order to test the significance of this, I calculated the difference between the self-bias and friend-bias score per participant and then ran an independent samples t-test on the resultant scores, comparing the choice vs. no choice groups. This test was significant ($t(190) = -3.05$, $p = .003$).

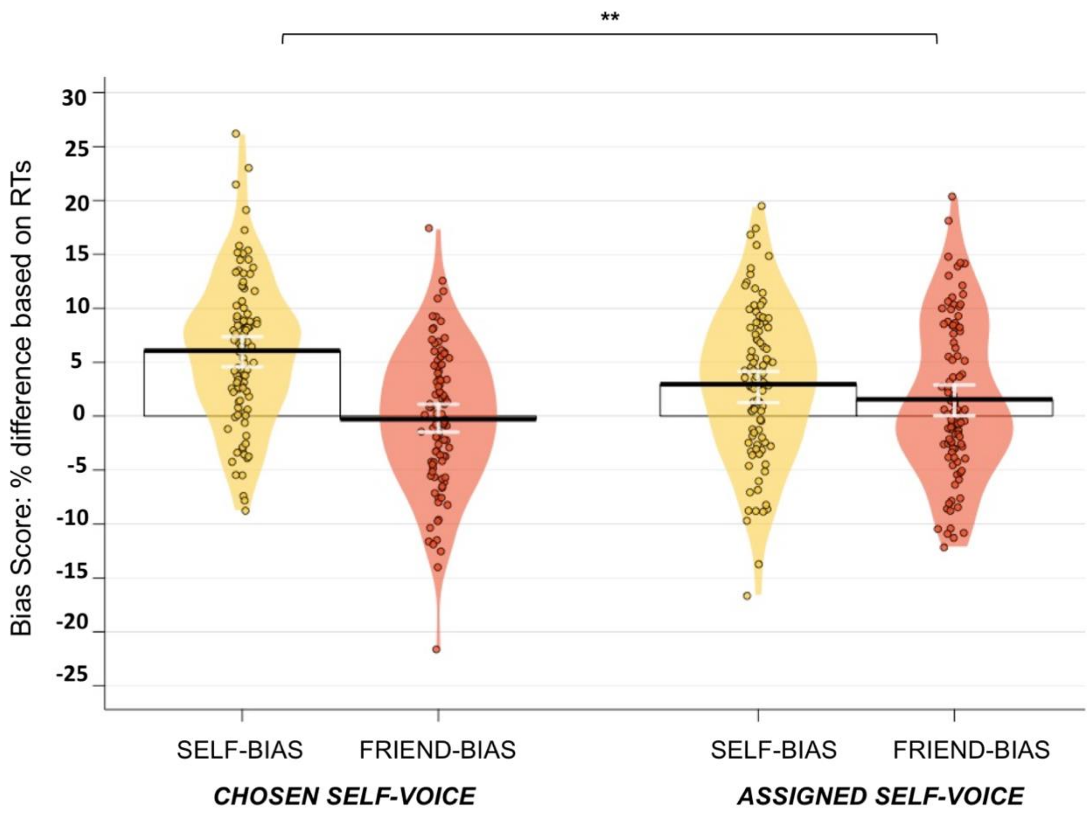

Figure 6. Exploratory analyses, Experiment 3: Measures of bias (self-bias vs. friend-bias) as a function of choice (choice vs no choice, i.e. whether the self-voice was personally chosen or randomly allocated) in mismatch trials only. The error bars indicate the SEs of the means. Coloured segments show smooth density curves for the full data distribution, while individual dots indicate mean scores per participant. *** p<.001

In Experiment 3, I asked whether the self-prioritisation effect could be modulated by choice. The results show that choice does interact with prioritisation: performance for the friend-voice is significantly reduced when it was *not* chosen as a self-voice. Thus, the self-voice is relatively more distinguished from the other voice identities. The results suggest that in rejecting certain voices as a self-voice, they are deemed less self-relevant and the social distance between the self and them is increased.

2.5 General Discussion

Across 3 experiments, I asked whether an unfamiliar, externally-generated voice could be incorporated into the self and, if so, which factors could modulate its self-relevance and subsequent bias.

In Experiment 1, I demonstrated that the concept of self can be quickly and flexibly extended to incorporate a new self-voice, as shown by quicker reaction times, higher accuracy, and increased perceptual sensitivity to the new self-voice relative to a friend or other's voice. This showed that it had been processed as a self-relevant stimulus, thus demonstrating a clear self-prioritisation effect. These results are a first demonstration that a voice does not need to be self-generated in order to become integrated into our concept of self and thus be prioritised in perception as a self-relevant stimulus. Given that voices need to be generated by an individual in order to exist (Scott & McGettigan, 2015), such an active signal should be intrinsically representative of another. Yet, this other-related social stimulus can be perceived as self-relevant by becoming associated with the self through ownership.

This finding is important when considering people who are unable to always use their own, self-generated voice to communicate. Specifically, the results support the notion that an alternative vocal signal could quickly become personally relevant as a means of self-expression.

In Experiment 2, I began to explore which factors might increase the bias afforded to a voice and, specifically, whether a voice would be made more self-relevant if it was gender-matched to the participant. However, I found no evidence to support this hypothesis; reaction times and perceptual sensitivity did not differ between gender-matched self-voices and non-gender-matched self-voices. Here then, gender-matching the new self-voice to the listener did not significantly moderate prioritisation of that voice. Instead, a significant self-prioritsation was evident in both groups for both RTs and accuracy.

Given how important voices are to representing our self (McGettigan, 2015), it might be surprising that a more personalised means of vocal expression did not generate a stronger bias, or greater prioritisation. In the memory literature, for instance, Cunningham et al. (2011) suggests that the strength of an item's encompassment within the self should, intuitively, be moderated by how far the item is autobiographically consistent and by how far it matches personal tastes and goals. Relatedly, Golubickis et al. (2019), have demonstrated that new stimuli are prioritised according to their identity relevance. Specifically, a stimulus that was associated with the self through a concept of high identity-relevance (i.e. musician) was more prioritised relative to a stimulus associated to a low identity-relevant concept (e.g. athlete). From this study, it might be predicted that a more identity-relevant stimulus – such as a gender-matched voice – would be prioritised more.

However, in Golubickis et al.'s (2019) study, it was not the identity-relevance of the *stimulus* itself (geometric shapes) that modulated prioritisation. Rather, it was the identity-relevance of the *concept* that the stimulus became associated with (musician, athlete). This is mirrored in the current study then; the relevance of the newly associated stimuli – the voices – did not modulate prioritisation. Instead, only the relevance of the concepts (self, friend, and other) influenced the degree to which the voices associated with those concepts were prioritised. As such, the voice associated with the self was relatively more prioritised than the friend-voice or other-voice as the concept of self was the most identity-relevant concept.

While it is possible that participants simply did not perceive a gender-matched voice to be more self-relevant, this null result gives rise to two further lines of enquiry. First, it must be considered that self-bias does not prioritise stimuli to different *degrees* according to their self-relevance. Rather, it may be a more binary system that determines whether a stimulus is, or is not, prioritised over others. In the current study, three unfamiliar voices became associated with the concepts of self, friend, and other. Within this task, the concept of self will always be the most identity relevant and so whichever stimulus is associated to that concept should be prioritised over the stimuli associated with either friend or other. If, however, participants were presented with two self-relevant options – for instance, a self-owned voice that was more identity-relevant (i.e., gender-matched) and a second self-owned voice that was relatively less identity-relevant (i.e., gender mismatched) – the more identity-relevant voice may be prioritised over the other in this instance. Here, however, I have assessed whether a self-owned voice is prioritised over a friend, or an other's voice *to a greater degree* because it is more identity-relevant and, indeed, it is not.

Second, it is possible that, in this context, it was irrelevant whether the new self-voice could more accurately represent the self or not. The perceptual matching task does not require social interaction; the new self-voice was never used as a primary means of sharing or representing the self-identity. Thus, it could be that a new voice which matches a participant's gender-identity may only be perceived as comparatively more self-relevant when that gender-identity is actively conveyed to others. Therefore, it may be that a communicative context and use of the voice to represent the self could modulate the bias afforded to it. This question is explored further in Chapter 6.

If bias over a new self-owned voice is not modulated according to how identity-relevant that voice is, Experiment 3 asked whether self-bias would be increased if the participant made a personal choice to take ownership of the voice. Specifically, I examined whether participants showed quicker reaction times and higher accuracy to the self-voice – relative to the friend-voice and other-voice – when the self-voice was chosen by the participant as compared to when it was randomly allocated.

The results showed that having a choice over the self-voice did interact with perceptual prioritisation. However, the shift in prioritisation did not manifest as quicker reaction times and increased perceptual sensitivity to the chosen self-voice relative to the unchosen self-voice. The interaction arose instead due to reduced performance for the friend-voice which had been rejected as a self-voice. In this instance, the friend-voice was processed as if it were an other's voice, with equivalent RTs for friend and other. Conversely, when the self-voice was randomly allocated and none of the voices were chosen nor rejected, there remained the typical distinction between the self-, friend- and other-voice.

While this indicates that choice increased the perceptual separation of the self-voice from the other identities, it seems to have happened via "othering" of the friend-voice rather than enhanced prioritisation of the self-voice per se. Interestingly, Huang and Wu (2016) have previously shown that the bias for self-owned items relative to other-owned items can be driven not only by more positive judgments for the self, but also by a negatively biased view towards the other-owned items. That is, in comparison to self-relevant stimuli, information relating to others can also be subject to derogation

(Huang and Wu, 2016). It is possible then that in rejecting the friend's voice as a possible self-voice, participants valued it relatively less and were negatively biased towards it. This could have led to its relative deprioritisation and, in turn, the increased distance between self and friend when the self-voice was chosen.

If this is the case though, it is unclear why the same 'other-derogation' and deprioritisation was not applied to the other's voice, which was also rejected as a potential self-voice. It is possible that processing of the other-associated stimulus in the perceptual matching task reflects the standard by which all non-self-relevant stimuli are perceived in the absence of self-bias. Against this baseline then, perception is relatively biased towards stimuli that are more self-relevant, including the friend's voice which is, socially, more self-relevant than an 'other'. However, more work may need to be done to fully elucidate how both choice and rejection in what becomes self-owned are valued, and how this shapes perception.

Furthermore, future studies may also benefit from exploring the interplay of ownership versus agency (i.e., choosing to take ownership) in this paradigm. Huang et al., (2009) have previously shown that people were biased towards items they chose to own relative to items they had been assigned ownership of. However, they were not biased towards items they had chosen for another participant to own. This suggests that stimuli may be perceived as especially self-relevant when they are both self-owned and self-chosen but that it is self-relevance itself that underpins the attribution of bias. Notwithstanding, the current study could be extended with a manipulation of whose voice is chosen, with some participants selecting the voice to represent themselves, and some choosing which voice should be assigned as the friend's voice.

In this initial suite of studies, I have begun to qualify which factors influence our perception of what is self-relevant or not and, thereby, which factors may aid the incorporation of a new, externally-generated voice within the self. I have shown that a new voice can be incorporated into the self after only brief exposure, as evidenced by a self-prioritisation effect in all three experiments. Moreover, personal choice in picking which voice to own may make it more self-relevant and allow that voice to be relatively more prioritised in comparison to voices that have been rejected. Finally, although matching the new self-voice to the gender-identity of the participant did not here modulate prioritisation of it, it is possible that the absence of an effect is due to the lack of a need to use this voice to represent oneself socially within the task. Given that the voice is a highly salient social stimulus, a more immersive, social interaction-based paradigm, in which the new self-voice is actively used to achieve communicated goals, could elicit a stronger bias than I have shown here. In such a context, prioritisation may be modulated by how closely the vocal stimulus can represent the self, including the gender identity of the speaker. This is explored further in Chapter 6.

Overall, the current results support the case for more personalised voices, at least to the extent that the speaker should have agency in choosing their new auditory identity. I have provided the first evidence that it is possible to incorporate a new voice into our concept of self, despite the fact that it should be perceived as belonging to an 'other'. Furthermore, this voice need not be self-generated in order to be deemed self-relevant. These studies thereby highlight the extent to which what is 'self' is flexible and can be adapted quickly to optimize perception towards what is immediately self-relevant. It remains unclear whether this new self-voice is incorporated into the physical representation of self sufficiently to alter perception of the true self-voice. However, on this data alone, it is evident that self can, at least temporarily, be dynamically extended to incorporate a new vocal identity.

3 Sense of agency over a new self-voice

Abstract

Within Chapter 3, two experiments explore sense of agency over a new, self-voice. Sense of agency is the feeling or belief that "I have caused that outcome". If a new voice has been incorporated into the self-concept through a sense of ownership over it, participants may also experience a greater sense of agency over it, relative to the voices of others. Sense of agency is typically measured by a perceived compression of time between a self-generated action and its outcome, in comparison to the time between a passively heard cue-tone and the same outcome. After an initial pilot study, Experiment 4 examined sense of agency over the self-voice relative to an other-voice when either voice was presented as the outcome of participants' actions. Results showed that participants perceived the time between their actions and the self-voice outcome to be relatively more compressed than when the outcome was the other-voice. This indicated a greater sense of agency over the self-voice. Experiment 5 furthered this by manipulating participant's engagement with postdictive cues regarding whether a voice belonged to self or other. Results replicated from Experiment 4 and further showed that sense of agency over a self-voice was increased when the relevance of the self-voice was emphasized. However, these results are tempered by the finding that a voice being self-owned or not similarly affected trials in which participants had no agency in generating an outcome. This may indicate that the effect of ownership over a voice interacts with temporal judgements outside of participants' sense of agency. Taken together with the previous findings of perceptual prioritisation of a new self-voice, these experiments expand the examination of integrating a new voice into the self-concept and how that voice may affect perpetual processing more widely thereafter.

3.1 Introduction

People behave with intentionality and thus have agency and control over their actions. Having agency is not only central to the self-concept but, more widely, to how self and other are differentiated and held accountable. Indeed, agency underpins the concept of responsibility and of self-awareness. Its centrality to our self-experience is illustrated by the fact that this agency is recognised by the criminal justice system (Haggard, 2017, p. 196); people typically act with intention. To navigate the world, it is essential that we understand the causality between our actions and their outcomes. This allows us to produce the *intended* outcome and to differentiate between the outcomes we ourselves have caused and those we have not. Relatedly, the experience or belief that "*I have caused that*" is known as our *sense of agency* (Gallagher, 2000).

Although Sense of Agency (SoA) has previously been measured by asking people to explicitly judge whether they caused an outcome or not, it is more commonly assessed by implicit measures using an intentional binding paradigm (Haggard, Clark, Kalogeras, 2002). According to Haggard (2017), people are rarely required to make explicit judgements of their agency in real life; an implicit measure – that captures the feeling of agency when no explicit evaluation of control is required – may therefore be more ecologically valid.

In an intentional binding paradigm (Haggard, Clark, Kalogeras, 2002) participants are asked to make temporal judgements about the times at which a) their voluntary, intentional, actions occur (i.e., a button press) and b) when the sensory outcome to that action occurs (i.e., an auditory tone). In the original intentional binding paradigm, participants judged the timings of these events by reporting where the hands were on a clock face when they occurred (see Figure 7). This measure of temporal judgement is called the Libet Clock method (Libet, Gleason, Wright, and Pearl, 1983). Importantly, participants were asked to judge the timings of actions and outcomes both when these events occurred sequentially – as causally related events – but also when either the action, or the outcome, occurred in isolation. The trials in which the events occurred in isolation were then used as baseline trials; temporal judgements could be compared against trials in which actions and outcomes both occurred.

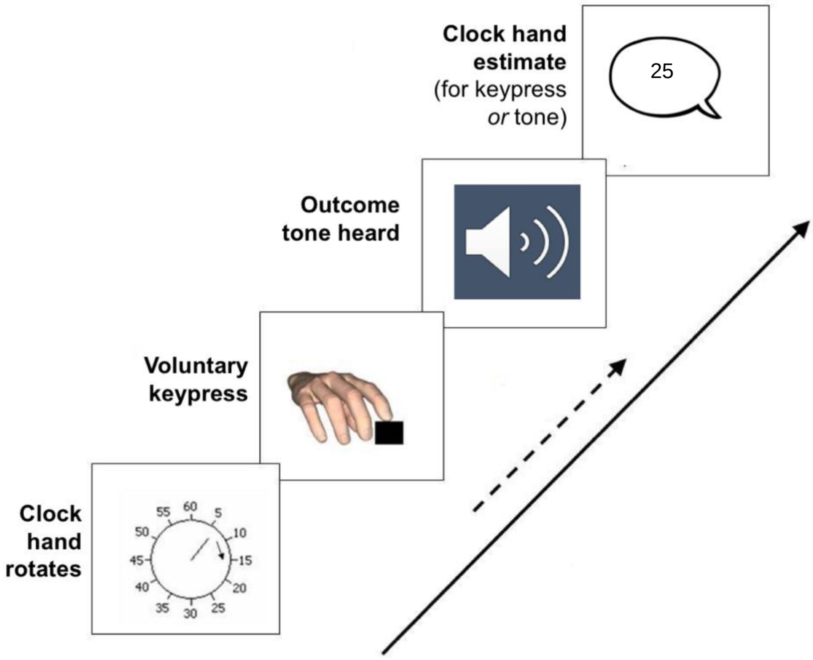

Figure 7. The typical trial structure in the active condition of a temporal binding paradigm using the Libet Clock method (i.e., Haggard, Clark, & Kalogeras, 2002). Participants make a voluntary keypress and, after a delay, hear a tone. The task is to estimate the position of the clock hand either when they pressed the key, or when the outcome tone sounded. Figure is reproduced from Moore & Obhi, 2012.

The results consistently revealed that participants perceived a compression of time between an action and its outcome. Specifically, participants perceived their action as occurring later than it actually did when a sensory outcome followed it (i.e. when the action didn't occur in isolation) and, moreover, the outcome was perceived as occurring earlier than it actually did when it was the result of self-generated action (i.e. the outcome followed an action) (Haggard et al., 2002). Thus, the time between a voluntary action and its sensory outcome was shown to be bound together and compressed perceptually. This perceived compression of time between a self-caused action and its outcome (Haggard, 2017) is widely established as an implicit measure of sense of agency.

The source of this compression of time, the binding of the two events, was theorised to be a result of the intentional nature of the action. Specifically, that critical sensorimotor events are bound together to allow us to experience a sense of agency over outcomes our voluntary actions have caused (Moore and Obhi, 2012). Indeed, the same effect of binding was absent or diminished for the outcomes of involuntary actions (i.e., induced via TMS; Haggard et al., 2002), suggesting that action-intention was key to binding. Hence, this temporal phenomenon became known as *intentional* binding (Haggard, Clark, Kalogeras, 2002).

The intentional binding effect – and the warped perception of time– has been robustly replicated since, using the original Libet clock method and by an interval estimation task (Engbert, Wohlschläger, Thomas, and Haggard, 2007). The latter task more simply asks participants to judge the duration of an interval between an action and its outcome rather than the timings of the two events separately, which lessens the cognitive load (Muth, Wirth and Kunde, 2021). Within the interval estimation task, sense of agency is typically derived from the perceived compression of time between a participant's voluntary action (i.e., keypress) and its subsequent outcome in an *active condition* (see Figure 8), relative to a *passive condition*. The passive condition comprises a cue-stimulus instead of an intentional starting action, that is then followed by an outcome. For instance, previous studies have used a forced involuntary action (Engbert et al., 2008), an auditory tone that is passively heard as a cue-stimulus (Imaizumi & Tanno, 2019), or a visual stimulus that is passively observed (Cravo, Claessens, & Baldo, 2009; Suzuki et al., 2019). In each instance, the passive cue-stimulus is followed by the same outcome as in the active trials. Thus, a binding effect is deemed present if interval estimates are significantly reduced in active trials – in which participants have agency in generating an outcome – and the passive trials, in which they do not.

For instance, in the seminal paper by Engbert et al. (2008) the authors manipulated both the agency (i.e., active, passive) as well as the agent (i.e., self, other), to assess the importance of the action being intentional and self-generated. As such, their study involved four conditions: an *active-self* condition in which participants voluntarily pressed a lever; a *passive-self* condition, in which participants were forced to press the lever involuntarily; an *active-other* condition, in which participants observed the voluntary action of another person pressing the lever and finally; a *passive-other* condition, in which participants observed another's forced action to press the lever. In all conditions, the outcome was an auditory tone played after an interval and the task was to verbally estimate the duration of the interval. Notably, their results revealed a significant interaction between agency and agent; there were shorter interval estimates in the active conditions relative to the passive conditions only for the conditions involving the self. Thus, intervals terminating in outcomes generated by the participants' own voluntary action (active self) were estimated to be significantly shorter than those intervals following an involuntary action (passive self). However, when only observing actions of others, interval estimates did not significantly differ between active and passive trials. This supported the notion that it is the intention to act that alters how the outcomes of those actions are perceived and thus gives rise to this perceptual compression of time. Moreover, that only under conditions of agency are modulations to people's sense of agency (such as self vs other) likely to become apparent.

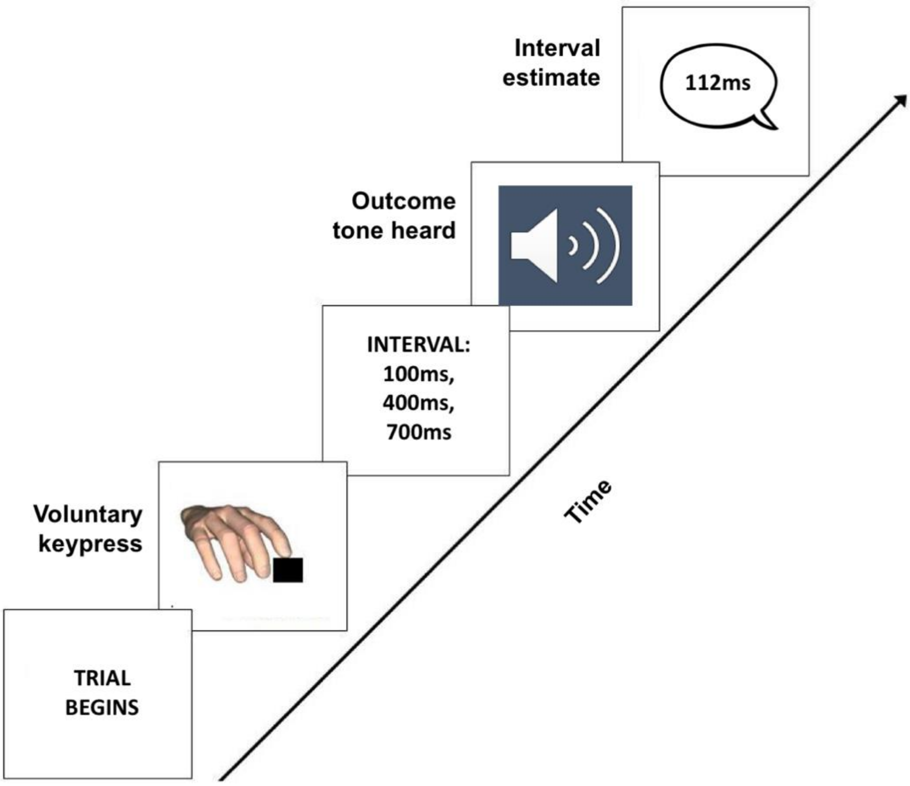

Figure 8. The typical trial structure in the active condition of the temporal binding paradigm using an interval estimation task (i.e., Engbert, et al., 2007). Participants make a voluntary keypress and, after an interval, hear a tone. The task is to estimate the duration of the interval. Figure is reproduced from Moore & Obhi, 2012.

Accordingly, several previous studies which have since explored *differences* in sense of agency (as opposed to whether participants experience a sense of agency or not) do not use a passive condition at all (Barlas & Kopp 2018; Obhi, Swiderski, & Brubacher, 2012; Ulloa, Vastano, George, & Brass, 2019; Zopf, Polito, & Moore, 2017). Instead, differences in sense of agency are assessed by an interval estimation task *in active trials only* according to the manipulation of interest. For instance, previous studies have required participants to make an action and then the congruency of the outcome has been manipulated (Zopf, Polito, & Moore, 2017), or the emotional valence of the outcome (Moreton, Callan, & Hughes, 2017), or Obhi, Swiderski, & Brubacher (2012) even manipulated how powerful people were primed to feel before making an action and experiencing the outcome. In such studies, a condition that gives rise to lower interval estimates would be said to show greater binding – and so greater agency – than a condition with higher interval estimates (Caspar, Desantis, Dienes, Cleeremans, & Haggard, 2016). Thus, it is *assumed* within such studies that when an action is made, people have a sense of agency over that outcome. The interval estimation task is then used to derive differences in strength of agency for actions and outcomes under certain conditions.

The variations in assessing sense of agency may partly underlie the mixed results that have arisen from such studies. For instance, it remains unclear whether it is action-intention, specifically, that drives the temporal compression. In the original study, the trials in which participants made an action (e.g., a keypress) were contrasted with trials where participants made involuntary actions that were induced by TMS (Haggard et al., 2002). In both conditions the outcome to these actions was an auditory tone. The presence of temporal binding between voluntary actions and the outcome – in comparison to involuntary actions and the outcome – led to the conclusion that the source of temporal binding was action intention. However, these two conditions differ by more than the lack of action intention. For instance, when participants are not in control of the action (i.e., in involuntary action trials), they also have less temporal control (Desantis et al., 2012; Hughes et al., 2013). This means the temporal onset of the subsequent outcome is less predictable. Therefore, the relatively reduced binding in this baseline condition may instead reflect the diminished temporal prediction rather than the absence of action intention (Pariyadath & Eagleman, 2007).

Further studies into this phenomenon cast growing doubt on action intention being the source of binding. Buehner (2012) highlights that sense of agency studies often conflate intention and causality. When generating the action intentionally, it may be easier to infer causality between that action and its outcome than when the action is involuntary or passively observed (Kirsch, Kunde, & Herbort, 2019). Therefore, it is possible that binding in the intentional action trials reflects only greater causal inference (Buehner, 2012; Eagleman & Holcombe, 2002). This argument is supported by the finding that temporal binding can even be shown between *involuntary* actions and their outcomes – that is, for action without intention – as long as self-causation has been implied (Dogge et al., 2012). Moreover, this binding is of equal magnitude to that demonstrated in earlier studies involving intentional action (Suzuki et al., 2019). Given this more recent evidence, this binding effect can also be referred to as *temporal* binding; allowing for the fact that action intention may not be as fundamental to the effect as first thought.

The finding that *implied* causation is sufficient to elicit binding is important because it also shows that our subjective experience can be modulated postdictively. Unlike the models proposing that it is action-intention – which relies on feed-forward mechanisms – that is the source of the temporal compression, this instead shows that it is possible for our perception and causal inference to be influenced by many different cues at different time points. This includes cues that alter our perception of time retrospectively (Stetson et al., 2006). The postdictive modulation of our sense of agency is further supported by a study by Takahata et al. (2012). Specifically, the authors manipulated the affective valence of outcomes and found that when it was a negative outcome (e.g., monetary loss), binding was attenuated. According to the authors, the fact that participants experienced a lesser sense of agency over actions which resulted in negative outcomes, highlights a self-serving bias in effect (Takahata et al., 2012) which postdictively modulates our perception. Notably, this explicitly shows that the degree of binding – and the sense of agency we experience – is intrinsically linked to the self-concept which pervasively biases our cognition.

Further to the original study then, it is evident that a range of cues can influence our subjective experience of time and the causality we infer between events. These cues include, but are not limited to: congruency between an action and its outcome and reliability of the outcome (Legaspi & Toyoizumi, 2019; Wolpe et al., 2013); inference of authorship (Ebert & Wegner, 2010); prior belief (Desantis, Roussel, and Waszak, 2011); efferent motoric information (Moore & Obhi, 2012); and the social context of the action (Pfister, Obhi, Rieger and Wenke, 2014). Overall, the above studies show that we experience a warping of time between our actions and their outcomes as long as we believe that we caused them.

Temporal binding studies typically examine agency between self-generated actions and the subsequent self-associated outcome. Thus, these studies rely on the implicit assumption that a self-generated action will cause a self-associated outcome precisely because it is self-generated. Interestingly though, binding can be postdictively modulated by prior beliefs regarding causality and self-bias (Desantis, Roussel, and Waszak, 2011), which suggests that an outcome being self-associated could postdictively imply that a self-generated action caused it. Thus, this paradigm could be inverted to measure whether certain outcomes are deemed self-associated *enough* to imply prior self-generated action. This offers an opportunity to test which sensory outcomes we can causally link to ourselves.

Here then, I ask whether an outcome that should not be possible for us to self-generate – the voice of another – can still be experienced as a sensory consequence of our action if the voice has previously become associated with the self through ownership. Earlier work presented in this thesis has demonstrated that we can incorporate a new voice into our concept of self and subsequently prioritise it in perception as a self-associated voice, even though it is the inherent biological property of another. In the current study, I explore whether temporal binding can be demonstrated between a self-generated action and a vocal outcome, if that outcome is a newly self-associated voice. To do so would not only demonstrate that we experience a sense of agency over that new voice but, ipso facto, that the new self-voice has been sufficiently incorporated into the self to be causally linked as a possible outcome of self-generated action. Such a result would therefore corroborate the previous findings presented in Chapter 2, that people can integrate a new auditory identity into their concept of self and, in turn, experience that new voice as ‘mine’ strongly enough to perceptually prioritise it and feel a sense of agency over it.

This rest of this chapter presents results from two experiments and associated pilot studies. First, I outline the procedure for the interval estimation paradigm that is used across all the experiments. Second, I detail pilot work that aimed to systematically measure how this procedure – specifically the differential trial structures in active trials (e.g., trials involving intentional action) and passive trials (e.g., those with no action) – can influence the degree of temporal binding. The results from the pilot influenced the task design of the two subsequent temporal binding studies that I then discuss. In Experiment 4, I asked whether there is an increased sense of agency over a new, self-voice using an interval estimation paradigm. Experiment 5 further examined temporal binding and the influence of self-association on the way we postdictively infer causality.

3.2 General Methods

3.2.1 Interval Estimation Task

Sense of agency is implicitly determined by measuring the perceived compression of time between two events: an action and its outcome. However, within the previous literature it remains inconsistent whether this compression of time between an action and its outcome is compared to a passive condition (with shorter interval estimates in the active condition being indicative of agency) or; is compared *within* the active condition relative to other actions and outcomes (with shorter interval estimates for any condition *within active trials* being indicative of a greater sense of agency). In light of these discrepancies, the experiments within this thesis derive a sense of agency by comparing a condition of agency (active condition) to one of non-agency (passive condition) according to the original paper by Haggard, Clark, & Kalogeras, (2002). Further, I employ an interval estimation task as in Engbert, Wohlschläger, Thomas, and Haggard (2007). This study by Engbert et al. provides the closest test case for the current study as the authors investigated both a condition of ‘agency’ (i.e., whether an action was voluntarily made or not: active or passive) and also a condition of ‘person’ (i.e. who made the action: self or other). Thus, within both active and passive trials, there was a further manipulation pertaining to ‘self or other’. In line with this, agency is here derived overall by shorter interval estimates in the active trials relative to the passive trials and *differences* in agency are derived according to trials featuring ‘self’ or ‘other’ within both active and passive trials.

In the studies within this thesis, the interval estimation task ran as follows (see Figure 9).

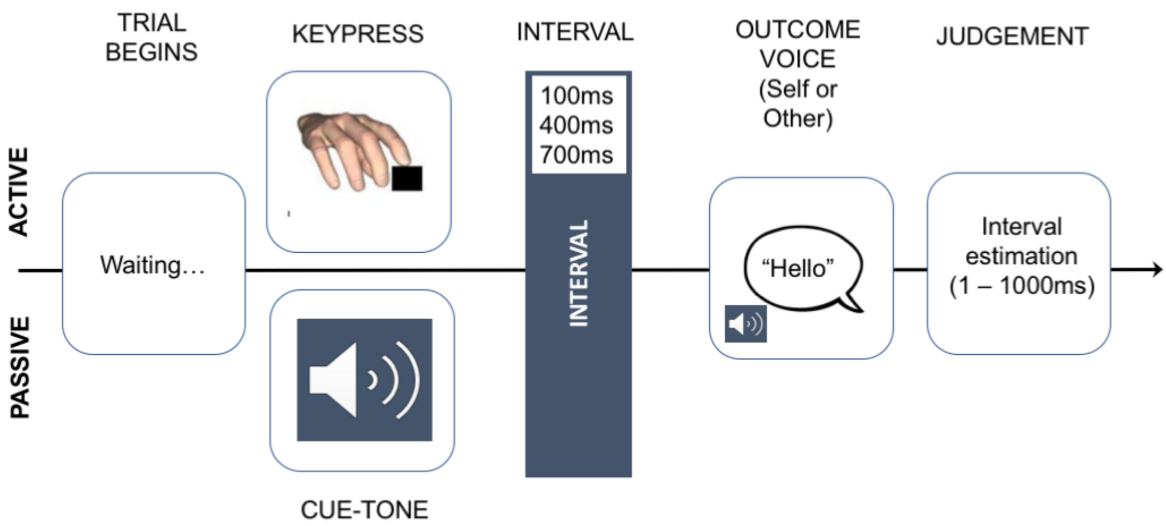

Figure 9. The trial structure used in the temporal binding paradigm within Experiment 4. The paradigm utilises an interval estimation task (i.e. Engbert, et al., 2007) involving both active and passive conditions (blocked). Trials in all conditions begin with a screen that says ‘Waiting…’. In active trials, participants then make a keypress prior to the interval. In passive trials, a 100ms cue-tone is heard prior to the interval. Participants hear either their self-voice or the other-voice after the interval. Participants judge the duration of the interval between either their voluntary keypress or the tone and the subsequent outcome (self-voice or other-voice).

In the active condition, the first screen said “Waiting…” in black letters in the centre of a white screen. The participant was instructed that whenever they saw this screen, they could press ENTER at any time of their choosing. When they pressed ENTER, the screen went blank and there was a period of silence – an interval – lasting either 100ms, 400ms, or 700ms. After the interval, an auditory stimulus played; a voice clip lasting ~500ms. This voice clip was the outcome of the button-press and was either the self-voice or an other’s voice. After the voice clip had ended, there was a further 500ms wait with a blank screen. The participant was then asked to estimate the duration of the interval period; the length of the silence between when they pressed

ENTER and when they heard the voice clip. Although this interval was always fixed to be one of three durations (100ms, 400ms, 700ms), participants were led to believe that this interval could be of any duration between 1-1000ms. Participants made their estimation response via a slider and clicked 'Next' to continue. The next trial automatically begins after this response and the initial "Waiting…" screen appeared again. No feedback was provided at any time.

In the passive condition, the "Waiting…" screen appeared as the first screen but there was no initial action from the participant – they did not press ENTER. Instead, the "Waiting…" screen remained visible for between 600-800ms with the exact period jittered across trials. This jitter was used to, as closely as possible, match the active condition in which participants pressed ENTER at a time of their choosing. According to a pilot study of 60 participants by Imaizumi & Tanno (2019), participants pressed the ENTER key after 703.6 ms on average (standard error of the mean 69.7). Thus, trial onset times of between 600-800ms mirrored the average time participants waited before voluntarily making the key-press. The "Waiting…" screen was ended by the onset of a cue tone instead that passively marked the offset of that screen and the onset of the interval. The interval was randomised to last either 100ms, 400ms, or 700ms, as in the active trials. After the interval, participants heard the same auditory stimulus as in the active trials – a voice clip lasting ~500ms. Again, the voice clip was either the self-voice or an other's voice. Participants were then asked to estimate the duration of the interval between hearing the tone and hearing a voice and made their response on the same slider as in active condition. No feedback was provided at any time.

There were four blocks: two blocks of active trials and two blocks of passive trials. The order of the blocks was randomised across participants. In each block, there were 30 trials (120 total) with each interval duration (100ms, 400ms and 700ms, with randomised presentation) presented an equal number of times in both active and passive conditions.

Finally, participants also completed 10 practice trials in which they were exposed to – and instructed through – both active and passive conditions. In these practice trials, participants did not hear a voice clip as the outcome of their actions but instead heard a click sound. Participants therefore had to estimate the time interval between key press and a click (active condition) or between a cue tone and a click (passive trials). The 10 practice trials each featured a different interval duration that ranged from 1-1000ms (i.e., the interval here was not fixed at either 100ms, 400ms, or 700ms), and participants received feedback on the actual interval duration after their estimation. After the interval estimation task, participants were debriefed and paid for their participation.

3.3 Pilot Study

It was critical for these the two agency conditions – active and passive – to be as closely matched as possible apart from the absence of action. This was to ensure that any difference in interval estimation between the two conditions could be derived as a difference in the sense of agency, and not because of the alternate task structure. In the procedure outlined above, the action in the active trials was a self-generated key-press and the equivalent cue in the passive trials was passive exposure to an auditory tone. This "active key-press or passive cue-tone" design replicates previous temporal binding studies. Importantly though, these studies use cue tones of different durations in the passive trials, specifically; 50ms (Pfister et al., 2014), 100ms (Imaizumi and Tanno, 2019) and 200ms (Desantis, Hughes, Wazsak, 2012). Here I examined whether these different cue-tone durations influence the perceived duration of the interval that follows it. If the interval duration is modulated by the cue tone duration, the difference in interval estimates for active and passive may not reflect only a difference in binding – and therefore of agency – but, rather, the difference in task structure across conditions.

The pilot study therefore manipulated the duration of the cue tone (50ms, 100ms, 200ms, and 400ms) and assessed the effect of these durations on subsequent interval estimates and, therefore, on temporal binding.

3.3.1 Participants

56 participants (age range = 19-40, mean age = 26.4 years, sd = 5.42 years, 28 female, 28 male) were recruited online via Prolific (www.prolific.ac). This data set

arose following the exclusion and replacement of 3 participants based on exclusion criteria outlined below. Participants were recruited as native speakers of English with no visual or hearing difficulties, and tested online using Gorilla (gorilla.sc, Anwyl- Irvine et al., 2019). Participants also had to have an approval rate of over 90% on the recruitment platform to be eligible, and were required to use Google Chrome as their internet browser. Finally, all participants had to pass a headphone check to ensure they were wearing headphones and were able to hear the stimuli.

Here, and in all further experiments within this thesis, a quicker and alternative headphone screening task was used in comparison to Experiments 1-3. Milne et al., (2020) provided implementation for a new headphone task based on the Huggins Pitch, an illusory pitch percept which can only be detected when stimuli are presented dichotically. Specifically, when wearing headphones, participants should be able to perceive a faint pitch within white noise due to a phase shift of 180° in only one channel. Thus, this pitch is not available when the noise is played over loudspeakers. In the task itself, participants heard three white noise stimuli and were asked to judge which one contained the hidden percept. As in the previous headphone task (Woods, Siegel, Traer, & McDermott, 2017) participants underwent six trials and only participants with 100% accuracy in detecting the hidden percept were eligible to continue. The screening task took less than 3 minutes to complete.

Ethical approval was obtained from the Departmental Ethics Committee in Speech, Hearing and Phonetic Sciences at UCL (SHaPS-2019-CM-030). Informed consent was obtained from all participants prior to testing and participants were paid and debriefed upon completion of the study.

3.3.2 Methods

3.3.2.1 Randomisation

Participants were randomised into 4 groups. Each group experienced a different cue tone duration within the passive trials: either 50ms, 100ms, 200ms, and 400ms. Note, the cue-tone duration is not the same as the interval durations, which were the same (100ms, 400ms, and 700ms) for all participants.

3.3.2.2 Design

The study consisted of two within-subjects factors “agency” (*active* vs *passive*) and “interval” (*100ms* vs *400ms* vs *700ms*), and one between-subjects factor “cue-tone duration” (*50ms, 100ms, 200ms,* and *400ms*).

3.3.2.3 Pre-processing

For each participant, I calculated a mean interval estimation in each condition e.g. Active trials with a 100ms interval were presented 10 times per participant and a mean estimate was calculated across these 10 trials. These mean estimations were used only as the basis of exclusion criteria (see below).

All trials which were more than 2 standard deviations above or below each participant’s mean (per condition) were eliminated, in line with previous studies (Engbert et al., 2008) as this demonstrated erratic performance. <2% of trials were eliminated on these criteria.

Finally, whole datasets of participants whose mean estimates did not increase monotonically with the presented interval duration (100ms, 400ms, 700ms) were also excluded and replaced as this demonstrates particular difficulty with the task and/or a lack of engagement (Suzuki et al., 2019; Caspar et al., 2016). 3 participants were excluded on these criteria by checking their mean estimates increased numerically with the interval estimate. Their data was replaced until a total of 56 participants (14 per cue duration) had successfully completed the study.

3.3.2.4 Analysis

I assessed whether there was a main effect of *cue-tone duration* on interval estimates in the passive trials. The interval estimates were assessed with linear mixed models (LMM) using lme4 (Bates et al., 2014) in the R environment (R Core Team, 2013). The model is detailed below:

E.g. lmer(interval estimates ~ cue tone duration + interval + 1 **|** participant + REML = "FALSE")

Statistical significance of the effects was established via likelihood ratio tests by dropping effects of interest from the appropriate model. For example, to establish whether there was a significant main effect of *cue tone duration,* this predictor was dropped from the model that included only the same random effects structure. For all analyses reported here, post-hoc comparisons were conducted in *emmeans* (Lenth, 2016) and were adjusted for the multiple comparisons via Bonferroni correction.

3.3.3 Results

Descriptive statistics for estimates at each interval according to the cue tone duration are given in Table 4 and plotted in Figure 10.

Table 4. *Pilot: Mean interval estimates (ms) according to the cue-tone duration.*

	Cue tone duration			
INTERVAL	50ms	100ms	200ms	400ms
100ms	137	151	201	185
400ms	336	373	425	400
700ms	553	599	694	611

Full model outputs to the LMM are reproduced in Supplemental Table 10. The main effect of *cue-tone duration* was significant ($\chi 2(3) = 9.18$, $p = .02$). Post-hoc comparisons showed that interval estimates differed significantly when the cue-tone duration was 50ms compared to 200ms ($p=.01$), with reduced estimates for shorter cue-tone durations. This demonstrated that the same time interval between the cue-tone and the outcome was perceived differently according to the duration of the cue tone itself (Figure 10). Indeed, the interval duration of 700ms was perceived as lasting for only 553ms on average when the prior cue-tone was 50ms but 694ms when the cue-tone was 200ms. This shows that participants' perception of time – and indeed the relative compression of time – may not only be reflective of agency but can also be affected by the duration of a proceeding auditory cue. Given that I cannot match active and passive trials perfectly – because the passive trials need to include a cue tone of at least some duration – it is therefore essential to exercise caution in interpreting the difference between interval estimates in active and passive trials in the experiments that follow. Rather, the difference in interval estimates *within* active trials and *within* passive trials may be more informative in deriving temporal binding.

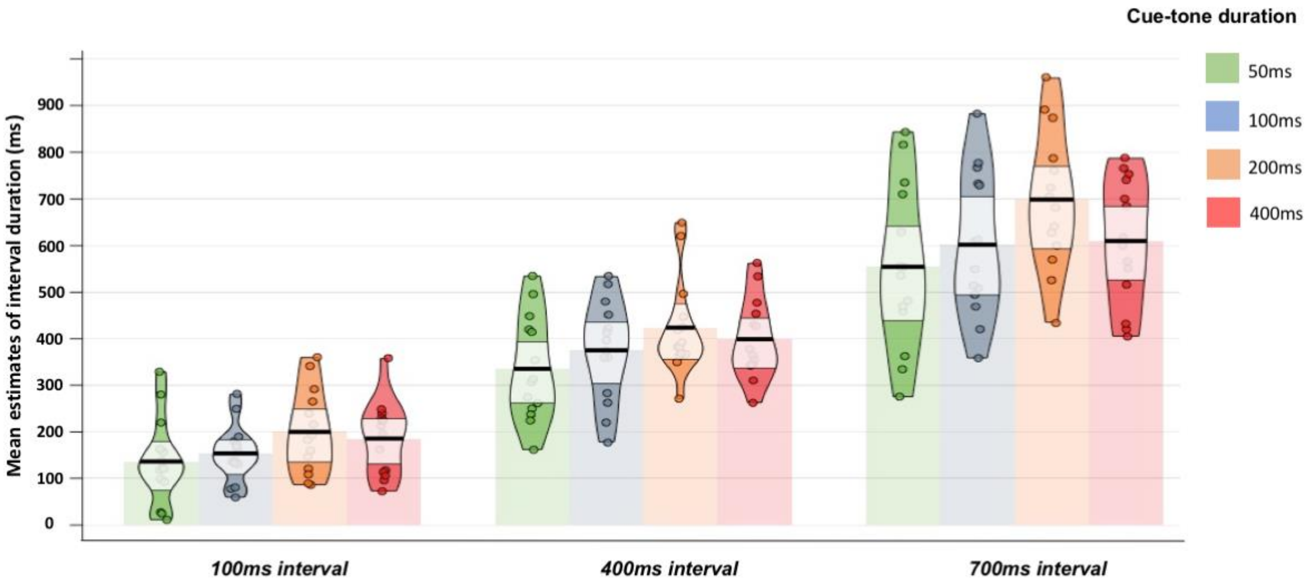

Figure 10. Pilot 1: Mean interval estimates (ms) according to duration of the cue-tone preceding the interval (50ms, 100ms, 200ms, 400ms). The inference bands indicate confidence intervals (CIs). Coloured segments show smooth density curves for the full data distribution, while individual dots indicate mean estimates per participant.

From this pilot study I have therefore established that the difference in interval estimates between active and passive trials may not reflect only a difference in agency but, also, the difference in task structure across conditions. Previous temporal binding studies have not addressed this issue; currently a temporal binding effect – and so a sense of agency – can be derived from the presence of significantly shorter interval estimates within the active condition relative to the passive condition (i.e., Imaizumi & Tanno, 2019; Engbert et al., 2008). In reporting the forthcoming temporal binding experiments in this thesis, I have taken steps to acknowledge the finding of my pilot study.

First, I use only one cue-tone duration of 100ms, which is the most frequently used duration across previous studies and is as closely matched to the active key-press as possible; the average duration of a key-press is 100-116ms (Dhakal et al., 2018). To directly compare with previous binding studies, I do still derive a main effect of agency from the difference between active and passive conditions. Indeed, a comparison of a condition of agency relative to a condition of non-agency is deemed a requirement of several sense of agency studies (i.e., Haggard et al., 2002; Engbert et al., 2008). Critically though, this interpretation is given in full acknowledgement that the difference between active and passive may not reflect *only* a difference in agency but may also be sensitive to other factors inherent in the difference in trial structure. Further work, that falls outside the remit of this thesis, will be critical to evaluating the temporal binding paradigm further and the extent to which factors affecting temporal perception, and so interval estimates, may have previously confounded measures of the sense of agency.

3.4 Experiment 4

3.4.1 Introduction

A voice cannot exist without prior action (Scott & McGettigan, 2015) and therefore voices are always associated with an agent. If we hear a voice, we expect it to have occurred via the action of either ourselves or another human. Indeed, voices are so strongly representative of an identity that for people who experience auditory verbal hallucinations, an imagined, individualised agent can be formulated in 80% of cases (Wilkinson & Bell, 2016).

Given that voices are so closely tied to agency, it is surprising that temporal binding studies have not examined our sense of agency in the context of voices more thoroughly. Currently, only two such studies exist. First, Ohata, Asai, Imaizumi and Imamizu (2021) examined the sense of agency over self-produced speech, specifically vowel sounds, and how this agency is attenuated by distorting the pitch of the auditory feedback. The authors ran an interval estimation task in which the participant's action was their production of speech and the outcome was the auditory feedback. This feedback (i.e., their voice) was either presented raw or distorted to be higher in pitch or lower in pitch than the original vowel sound. This outcome was presented after an interval of either 200ms, 400ms, or 600ms. Shorter interval estimates (e.g., stronger binding) were demonstrated for the unaltered feedback relative to the distorted feedback; raising or lowering the pitch caused an attenuation in binding and a lessening of the sense of agency over that outcome. This study shows that we have a sense of agency over our voice actions and the agency we experience is sensitive to cues which lessen the causal link between our action and its outcome. As mentioned above, these cues include congruency between an action and its outcome – which has been manipulated here. The altered feedback sounded less 'like self' than expected and the sense of agency over that outcome was therefore reduced.

In contrast to this study, Limerick, Moore & Coyle (2015) demonstrated that there was a lesser sense of agency over outcomes that have been generated via a voice command compared to outcomes generated via a button press. Specifically, participants either pressed ENTER, experienced a 500ms interval and then heard a tone *or* participants said, "Go" and then experienced a 500ms interval and heard a tone. Temporal binding was here measured with the original Libet clock method (e.g., Haggard et al., 2002) such that participants estimated the interval between their actions (saying "Go" or a button press) and an outcome (a tone) both when they occurred sequentially and in isolation. Strikingly, the authors found a significant effect of temporal binding for the key-press condition but not for the voice command condition.

This study suggests that people have less agency over outcomes they generate with their own voice than we might, intuitively, expect. It should be highlighted that this study differs to that of Ohata et al., (2021) in that the voice is here the action rather than the outcome. Problematically, in the voice command condition in this study, a voice recognition interface processed the voice and began the interval at the offset of the voice command, rather than the onset. The authors acknowledged, however, that participants perceived their action as actually occurring at the beginning of their voice command, rather than at the end which would, in effect, lengthen the interval between action and outcome. Further, the authors also report that the average time to produce the utterance "Go" takes ~300ms. It is possible therefore that the considerably longer time it takes to produce this voice command relative to a key-press (~100ms; Dhakal et al., 2018) affects the perception of time in the following interval. Indeed, in the earlier pilot study, longer cue-tone durations resulted in participants perceiving longer interval durations which, in Limerick et al.'s voice command condition, may have been enough to diminish any apparent temporal binding.

Intuitively, one would expect to see evidence of a sense of agency over our voice because we are the agents of our voice. Further, our voice is inherently self-associated and, always, the outcome of our self-generated action. The lack of agency in the study by Limerick, Moore & Coyle (2015) could be because the outcome was a tone, both when the action was a voice command and when the action was a keypress. That is, in both instances, the outcome tone had been generated by self-action but it was no more self-associated in one condition than the other. In contrast, in Ohata et al.'s study (2021) the outcome was the participant's own voice and so was strongly self-associated. Here, the authors did show temporal binding and, therefore, a sense of agency. It is important then to examine the identity of the outcome and, particularly, the specific influence of a self-associated outcome on temporal binding. Previous studies implicitly assume the outcome will be self-associated if it has been self-generated and, thus, most studies do not consider the specific influence of the outcome depending on whether it is self-associated or not. To examine the self-associated nature of the outcome can further our understanding of the sense of agency.

Makwana & Srinivasan (2019) has previously examined this in part, by first using a perceptual matching paradigm (see Chapter 1 for full procedure) to train new associations regarding the outcome. Specifically, participants were asked to associate different identities – self, friend, and stranger – each with a different geometrical shape –circle, square, and triangle (e.g., Sui et al., 2012). Thereafter, participants completed an interval estimation task: participants made a key press (the action) and experienced an interval (either 100ms, 400ms, or 700ms), after which either a circle, a square, or a triangle (the outcome) was presented visually on screen. Participants estimated the time between their key-press and the outcome and did not complete a passive condition. Thus, sense of agency was derived from which of the outcomes (self-, friend-, other-associated outcomes) gave rise to shorter interval estimations. The results showed that participants perceived the interval between their action and the outcome to be shorter when the outcome was the shape associated with the self, relative to the shapes associated with either friend or stranger. This showed that self-association can influence our sense of agency and, crucially, enhance it.

The present study builds on this by examining sense of agency over a voice that, importantly, is not inherently our own but has become self-associated. Here, I asked whether there would be a greater sense of agency over a new self-owned voice relative to voices associated with others. This relies on the new self-voice having been sufficiently incorporated into the self-concept to be causally linked as a possible outcome of self-generated action.

3.4.2 Hypothesis

I hypothesised that participants would have a greater sense of agency over a newly self-associated voice relative to a voice associated with an other. Given that participants only have agency in generating outcomes in the active trials, modulations to their sense of agency should only be apparent within these active trials. Indeed, only the active trials are a condition of agency whereas passive trials are a condition of non-agency. Thus, if the voice identity of the outcome (self-voice, or other-voice) does influence participants' sense of agency, this would be indicated by an interaction between *voice identity* and *agency*, with shorter estimates for intervals terminating in the self-voice relative to the other-voice, in the active trials only.

3.4.3 Participants

42 participants (mean age = 27.5 years, SD = 5.66 years, age range = 18–40, 21 female, 21 male) took part in the study. This data set arose following the exclusion and replacement of 5 participants based on exclusion criteria outlined below. A minimum sample size of N=41 was determined by an a priori computation in G*Power (3.1.3, Faul et al. 2007) in which a medium effect size of 0.6 could be detected in a repeated measures ANOVA with $\alpha = .05$ and power of $1 - \beta = .90$. This sample size is in line with previous studies (e.g. Kirsch et al., 2019; Ruess, Thomaschke, & Kiesel, 2017) in which is it assumed that intentional binding is a medium-effect (d=0.593). Due to within-task counterbalancing of voice stimuli, a final sample size 42 participants was needed to allow for an equal number of participants (N=7) to complete the task under each counterbalanced route.

All participants were recruited online via Prolific (www.prolific.ac) as native speakers of English with no visual impairments or hearing difficulties, with over 90% approval rate on the platform. Participants were tested online using Gorilla (gorilla.sc, Anwyl-Irvine et al., 2019). Participants were also required to use Google Chrome as their internet browser and to pass a headphone check (Milne et al., 2020) to ensure they were wearing headphones and able to hear the stimuli.

None of the participants had taken part in any of the pilot studies associated with this project and, upon completion of the study, were paid for their participation. Ethical approval was obtained from the Departmental Ethics Committee in Speech, Hearing and Phonetic Sciences at UCL (SHaPS-2019-CM-030), and informed consent was obtained from all participants prior to testing.

3.4.4 Methods

3.4.4.2 Stimuli

The voice stimuli were the same as those used in Chapter 2. Specifically, they were auditory exemplars of three male speakers and three female speakers each saying ‘hello’ in Southern Standard British English. All voice stimuli had been rated in an earlier pilot and found to be well-matched in social attractiveness, trustworthiness, and discriminability. They were further normed for amplitude via RMS-norming (see Chapter 2 for full details).

3.4.4.3 Procedure

The experiment involved three tasks; the first asked participants to choose a new voice, the second ensured recognition of that voice compared to others, and the third tested the sense of agency over the voice.

Task 1: Choosing a new voice

Participants were invited to choose 1 of 2 voices to be their new voice and to represent them in the rest of the study. Both available voices were gender-matched to the identity of the participant, and participants could play a clip from both voices before choosing. For each participant, the two voices made available for selection from the total set of six available voices were determined first by the participant’s gender (voices were gender-matched) and then by counterbalancing, ensuring that each possible pairing of voices was available for selection by an equal number of participants across the experiment. After selection, whichever voice had been rejected by the participant was discarded and did not then feature in the later tasks.

Task 2: Voice recognition

Before testing the sense of agency participants had over a new self-voice, it was important to test whether participants could recognise that voice and distinguish it from the voice of an ‘other’. I therefore ran an adapted version of a perceptual matching task (Sui et al., 2012) to train the two new associations: 1) between themselves and their new voice, and 2) between an ‘other’ and the other’s new voice.

The task started with passive exposure to each of the two new voices and their associated identity label, YOU or STRANGER. Following a 500ms fixation cross, an identity label was displayed in the centre of the screen in black uppercase font on a white background and remained on-screen for 3000ms. 500ms after the label’s onset, an auditory exemplar of ‘hello’ from the correct voice for that label was played, lasting approximately 500ms. After the auditory stimulus finished, the label remained on the screen until the end of the 3000ms trial. Stimuli were presented in a random order. This familiarisation phase consisted of 8 trials, with each label-voice pairing being presented four times. This phase thus lasted approximately 1 minute, including on-screen instructions, and was immediately followed by the test phase.

In the test phase, each trial started with a 500ms fixation cross in the centre of the screen, after which an auditory exemplar played for its total duration (approx. 500ms). Immediately after the auditory offset, the word ‘YOU’, or ‘STRANGER’ was displayed in the centre of the screen. Participants were asked to judge whether the identity label on-screen was a match or mismatch to the voice heard and to respond by keyboard press as quickly and as accurately as possible. Participants were instructed to press the left arrow for ‘MATCH’ and the right arrow for ‘MISMATCH’. This left-right ordering of match and mismatch remained constant on all trials, for all participants. Feedback was given on-screen for 500ms immediately following every response: a green tick for correct, a red cross for incorrect, and text feedback of ‘TOO SLOW’ for responses occurring after 1500ms. The next trial began after the 500ms feedback period was complete.

Participants performed only one block of 40 trials, in which the order of match vs mismatch trials was randomized for all participants. Participants were informed of their overall accuracy at the end of the block. All participants continued to complete both tasks, but accuracy within this voice recognition task was later used as the basis of an exclusion criterion (see below).

Task 3: Interval Estimation Paradigm

The procedure for the interval estimation was the same as in the earlier pilot study (see section 3.2.1 for details) with a few key differences:

As before, in the active condition, the first screen displayed “Waiting…” and the participant was instructed to press ENTER at any time of their choosing. When they pressed ENTER, there was an interval lasting either 100ms, 400ms, or 700ms. Critically, after the interval, one of two possible auditory stimuli played: either a clip of the *voice* that they had chosen as a self-voice or a voice they had been told belonged to an “other”. Similarly, in the passive condition, the “Waiting…” screen appeared until a 100ms tone passively marked the onset of the interval, lasting either 100ms, 400ms, or 700ms as in the active trials. After the interval, participants heard either the self-chosen voice that belonged to them or the voice that belonged to the other. I selected a cue-tone duration of 100ms in the passive trials based on the outcomes of earlier pilot work.

There were four blocks: two blocks of active trials and two blocks of passive trials. The order of the blocks was randomised across participants. In each block, there were 30 trials (120 total) such that each voice (self, other) was presented 10 times after each interval duration (100ms, 400ms and 700ms, with randomised presentation), in both active and passive conditions. After the interval estimation task, participants were debriefed and paid for the participation. The whole study took participants 35 minutes to complete, on average.

3.4.4.4 Design

The experiment consisted of three within-subject factors: “agency” (*active* vs *passive*), “voice Identity” (*self* vs *other*) and “interval” (*100ms* vs *400ms* vs *700ms*).

3.4.4.5 Pre-processing

For each participant, I calculated a mean interval estimation per participant in each condition e.g., Active trials with a 100ms interval were presented 10 times for each outcome (self-voice, other-voice) per participant and a mean estimate was calculated across these 10 trials. These mean estimations were used only to apply the exclusion criteria (see below).

3.4.4.6 Exclusion Criteria

It was important to ensure that every participant was sufficiently familiar with the voices assigned to the self and to the other before participants' the sense of agency over each voice was tested. Task 2 therefore measured participants' recognition accuracy. Participants whose performance accuracy was at or below chance (≤50% + 95% CI) were excluded and replaced as this demonstrated a lack of engagement and/or an inability to distinguish between the self-voice and the other's voice. Three participants were excluded on this criterion and their data replaced.

In the interval estimation task, whole datasets of participants whose mean estimates did not increase monotonically with the presented interval duration (100ms, 400ms, 700ms) were also excluded as this demonstrated particularly difficulty with the task and/or a lack of engagement (Suzuki et al., 2019; Caspar et al., 2016). A further 2 participants were excluded on this criterion by checking their mean interval estimates increased numerically with the interval duration. Their data was replaced until a total of 42 participants had successfully completed the study.

Lastly, all trials which were more than 2 standard deviations above or below each participant's mean (per condition) were eliminated in line with previous studies (Engbert et al., 2008) as this demonstrated erratic performance. 2.8% of trials were eliminated on these criteria.

3.4.4.7 Analysis

Voice Recognition Task

In the voice recognition task, I measured trialwise accuracy per participant for both the self-voice and the other-voice. Participants were able to perform the task with high accuracy (84%), showing that they were sufficiently able to distinguish between the self-voice and the other-voice.

Interval Estimation Task

Interval estimations were assessed with linear mixed models (LMM) using lme4 (Bates et al., 2014) in the R environment (R Core Team, 2013). I ran a model that included an interaction between *agency* and *voice identity* to first establish whether there were significant differences in interval estimates in passive and active trials according to the voice outcome: self or other. This model thus included an interaction between *agency* and *voice identity,* fixed effects of *agency* (active, passive), *voice* (self, other) and random intercepts of *participant* and *interval.* 'Interval' was included as a random intercept as it was not a factor of interest to the main analysis. Given that all participants whose interval estimates did not increase monotonically with the interval duration were excluded, it was expected that the effect of interval would be significant in every analysis. However, this would merely reflect the fact that interval estimates did significantly differ according to the true interval duration of 100ms, 400ms, or 700ms, as expected. Thus, 'Interval' was included as a random intercept within models in all further analyses across the thesis.

E.g. lmer(interval estimates ~ agency * voice identity + 1 | participant +
1 | interval, REML = FALSE)

Statistical significance of the effects was established via likelihood ratio tests by dropping effects of interest from the appropriate model. For example, to establish whether the interaction was significant, the interaction was dropped from a model including the other fixed effects and random intercepts. For all analyses reported here, post-hoc comparisons were conducted in emmeans (Lenth, 2016) and were adjusted for multiple comparisons via Bonferroni correction.

For LMMs, the models' estimates and associated confidence intervals for each effect are reported. The further away from 0 the estimates are, the bigger the effect. If the confidence intervals do not cross 0, the relevant effect is significant.

3.4.5 Results

Descriptive statistics for interval estimates (ms) for both active and passive trials, at each interval according to the voice outcome are given in Table 5.

Table 5. *Mean interval estimates (ms) across conditions in Experiment 4.*

	Active Trials		**Passive Trials**	
INTERVAL	Self-voice	Other-voice	Self-voice	Other-voice
100ms	132	144	159	184
400ms	311	314	363	365
700ms	509	527	580	594

To assess sense of agency, I ran an LMM on interval estimates to determine whether there was evidence of temporal binding and, further, whether this binding was modulated according to the voice identity of the outcome (self, or other). The interval estimates are plotted in Figure 11a and 11b and full model outputs from the LMM are reproduced in Supplemental Table 11.

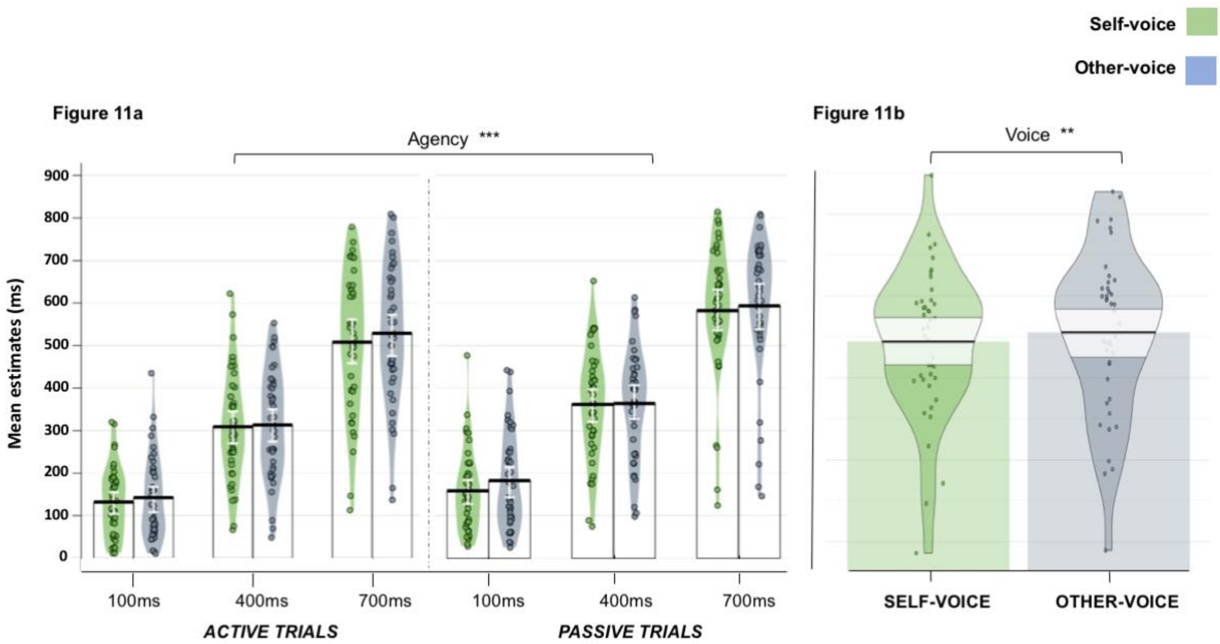

Figure 11a. Experiment 4**:** Mean interval estimates (ms) according to the outcome (self-voice or other-voice), the true interval duration (100ms, 400ms, and 700ms), in both active and passive trials. The error bars indicate the SEs of the means. Coloured segments show smooth density curves for the full data distribution, while individual dots indicate mean estimates per participant. **Figure 11b.** Experiment 4**:** Mean interval estimates (ms) according to the outcome (self-voice or other-voice), collapsed across all conditions. Inference bands indicate CIs. Horizontal bars show main effects with asterisks denoting significance within estimates as determined by likelihood ratio tests. *** p<.001

The interaction between *agency* (active or passive) and *voice identity* (self, or other) was non-significant. Rather, there was a significant main effect of *agency* ($\chi2(1)$ = 121.23, p = <.001) showing that interval estimates were significantly lower in the active trials than the passive trials (p<.001) across both self- and other-voice outcomes. Thus, in the trials in which participants had agency in generating the voice outcome, interval estimates were significantly lower. This result is on par with previous studies and this effect of temporal binding is typically interpreted as an implicit measure of sense of agency. However, as shown in the earlier pilot study, the difference between estimates for active and passive trials may also be affected by differences in trial structure. Thus, it cannot be concluded that this effect solely represents differences in the sense of agency.

Importantly though, there was a significant main effect of *voice identity* ($\chi2(1)$ = 6.837, p= .008), showing that interval estimates were systematically lower for intervals ending in the self-voice than the other-voice ($p=.008$). That is, participants experienced the interval preceding the outcome as shorter if the outcome was the self-voice. This suggests that participants' temporal judgements were affected by the presence of a self-associated outcome. However, the fact that the main effect of *voice identity* was also present in the passive condition – in which participants did not have agency in generating the outcome – requires closer examination as discussed below.

3.4.6 Discussion

In the current experiment, there were significantly shorter interval estimates in the active trials relative to the passive trials which, according to previous literature, shows a temporal binding effect and a sense of agency. Thus, the paradigm as implemented here has successfully replicated previous findings. It should still be noted though, that the difference between active and passive trials may also reflect the differences in task structure as suggested by the pilot study.

Of key interest here is the fact that there was also a main effect of *voice identity*. That is, when the interval terminated in the self-voice relative to the other-voice, it was perceived as being briefer in duration. In the active trials, participants had equivalent agency in producing both outcomes: they made the same action (a keypress) to generate either voice. Thus, to generate either voice required the same level of temporal control and the same level of action intention, and these internal cues are widely reported to influence sense of agency (Haggard et al., 2002; Moore and Obhi, 2012). The difference in estimates for intervals terminating in the self-voice relative to the other-voice, must therefore be by virtue of the outcome itself. Indeed, external cues as to the identity of an outcome (i.e., as belonging either to self or other) modulated participants' perception of interval durations. Importantly then, this influence must be postdictive, suggesting that peoples' sense of agency can be retrospectively adjusted according to the outcome experienced. It is possible that further cues pertaining to the outcome's identity – as either self, or other owned – enabled participants to derive a greater degree of causality between their actions and their outcome when the outcome was self-owned.

This result tallies with a previous study by Makwana and Srinivasan, (2019), in which binding was greater when actions resulted in a self-relevant outcome in comparison to actions that resulted in an outcome associated with an other. This finding is key because it calls into question the factor *driving* greater binding and so, sense of agency. Typically, greater binding is posited to arise because outcomes have been self-generated (Chambon & Haggard, 2012) and/or because their self-generation allows for stronger causal belief between the action and its outcome (Desantis et al., 2011; Haering & Kiesel, 2012). However, these results along with Makwana and Srinivasan's (2019) findings suggest that temporal binding may also be driven by the fact that the outcomes are self-relevant. This has been implicitly assumed in previous studies as a sense of agency is typically interpreted in conditions when outcomes have been self-produced as compared to outcomes that have not been self-produced. Yet, inherently, in self-producing an outcome, that outcome will become self-relevant. In the current study, the status of an outcome as being self-relevant was partially separated from whether it had been self-produced or not; all outcomes were self-produced, but one outcome was more self-relevant (through ownership) than the other. This suggests that the perceived self-relevance of an outcome could play a key role in participant's sense of agency.

In the first instance then, Experiment 4 supports the hypothesis that, in owning a new self-voice, people experience a greater sense of agency over that voice. Critically though, this result may be tempered by the fact that the effect of *voice identity* was also present in the passive trials. Indeed, there was not an interaction effect as predicted by the results of Engbert et al., (2008). This is surprising given that the passive trials involved no action and should therefore not reflect a difference in sense of agency for each voice. Clearly, self-association influences interval estimates but it must be considered that this influence may be unrelated to a sense of agency if it also occurs in conditions of no agency.

It remains unclear then why the self-associated nature of an outcome affects temporal judgements in the passive trials, a condition of non-agency. Thus, two related questions arise from Experiment 4. First, it is interesting that the self-relevance of an outcome may affect participants' temporal judgements. To further probe this, Experiment 5 aimed to increase the relevance of the voice outcome's ownership (as being either self-owned or other-owned) by additionally including a postdictive judgement as to whether the outcome was the self-voice or the other-voice. Second, I wanted to rerun the study to explore whether the significant effect of *voice identity* on temporal binding in the passive trials would replicate.

3.5 Experiment 5

3.5.1 Hypothesis

I hypothesised that participants would have a greater sense of agency over a newly self-associated voice relative to a voice associated with another. This result would be indicated via shorter interval estimates for the action-outcome trials in which the self-voice is the outcome compared to trials in which the outcome is the other-voice. If self-ownership over a voice affects participants' *sense of agency* – rather than temporal judgement more broadly – this should be indicated by an interaction between *voice identity* and *agency*, such that the effect of *voice identity* is present in the active trials only. Conversely, if self-ownership of a voice affects participants' temporal perception more widely – and possibly outside of notions of agency – this should be indicated by the effect of *voice identity* also being present in the passive trials.

3.5.2 Participants

42 participants (mean age = 28.4 years, SD = 5.83 years, age range = 18–40, 21 female, 21 male) took part in the study. This data set arose following the exclusion and replacement of 7 participants based on exclusion criteria outlined below.

3.5.3 Methods

3.5.3.2 Stimuli

The voice stimuli were, again, the same as those used in Chapter 2 and here in Experiment 4. Specifically, they were auditory exemplars of three male speakers and three female speakers each saying 'hello' in Southern Standard British English.

3.5.3.3 Procedure

Experiment 5 replicated the procedure of Experiment 4 apart from the addition of a postdictive question the voice's ownership in the interval estimation task (see Figure 12 below).

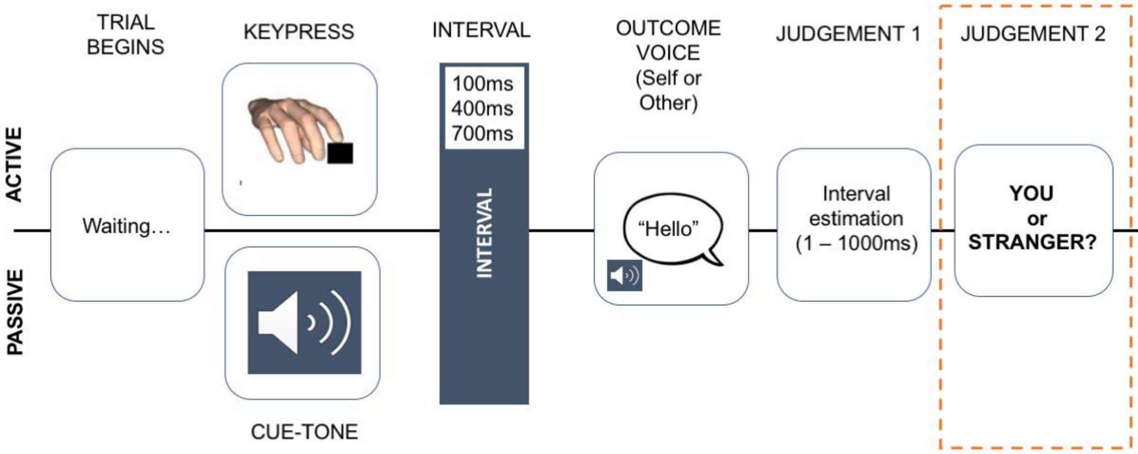

Figure 12. The trial structure used in the temporal binding paradigm within Experiments 5-7 and 10. The paradigm utilises an interval estimation task (i.e., Engbert et al., 2007) involving both active and passive conditions (blocked). Trials in all conditions begin with a screen that says 'Waiting…". In active trials, participants then make a keypress prior to the interval. In passive trials, a 100ms cue-tone is heard prior to the interval. Participants hear either their self-voice or the other-voice after the interval. Participants judge the duration of the interval between either their voluntary keypress or the tone and the subsequent outcome (self-voice or other-voice). Newly, participants also judge which identity self, or other) the voice outcome belonged to.

Specifically, participants experienced the same trial structure as before in both the active and passive condition: a key-press/auditory tone, an interval, and then a voice clip of either the self-associated voice or the other-associated voice. Participants then made their interval estimation response, as before. However, following this, participants were additionally asked "*Who did the voice belong to? YOU or STRANGER"* and gave their answer via an on-screen click. The labels for the two voices were chosen as they align with the voice-identity labels in the previous voice recognition task. After participants had made their response, the next trial automatically began and the initial "Waiting…" screen appeared again to begin another trial.

3.5.3.4 Design

The experiment, as before, consisted of three within-subject factors: ‘agency’ (*active* vs *passive*), ‘voice’ (*self* vs *other*) and ‘interval’ (*100ms* vs *400ms* vs *700ms*).

3.5.3.5 Pre-processing

As in Experiment 4, I calculated a mean interval estimation per participant in each condition e.g., Active trials with a 100ms interval were presented 10 times for each outcome (self-voice, other-voice) per participant and a mean estimate was calculated across these 10 trials. This mean estimation was used only to apply the exclusion criteria (see below).

Further, a trialwise measure of accuracy was generated for participants’ responses on: “*Who did the voice belong to? YOU or STRANGER”.*

3.5.3.6 Exclusion Criteria

Whole datasets for participants whose performance in the voice recognition task was at (or below) chance level (≤50% + 95% CI) were be excluded and replaced as this demonstrated a lack of engagement and/or an inability to distinguish between the self-voice and the other’s voice. 4 participants were excluded on these criteria and their data replaced.

In the interval estimation task, whole datasets of participants whose mean estimates did not increase monotonically with the presented interval duration (100ms, 400ms, 700ms) were also excluded and replaced as this demonstrated particularly difficulty with the task and/or a lack of engagement (Suzuki et al., 2019; Caspar et al., 2016). A further 3 participants were excluded on these criteria and their data replaced until a total of 42 participants had successfully completed the study.

Further, all trials which were more than 2 standard deviations above or below each participant’s mean (per condition) were eliminated in line with previous studies (Engbert et al., 2008) as this demonstrated erratic performance. 2.7% of trials were eliminated on these criteria.

Lastly, given the addition of the postdictive question: “Who did the voice belong to?”, I did not analyse erroneous responses. Trials in which the participant had incorrectly identified the voice were eliminated. This was because incorrect recognition of the voice could affect the inference of causality and therefore the sense of agency; 8.1% of trials were eliminated on these criteria.

3.5.3.7 Analysis

Voice Recognition Task

In the voice recognition task, I measured trialwise accuracy per participant for both the self-voice and the other-voice. Participants were able to perform the task with high accuracy (87%), showing that they were sufficiently able to distinguish between the self-voice and the other-voice.

Interval Estimation Task

As in Experiment 4, interval estimations were assessed with linear mixed models (LMM) using lme4 (Bates et al., 2014) in the R environment (R Core Team, 2013). I ran a model that included an interaction between *agency* and *voice:*

e.g. lmer(interval estimates ~ agency * voice identity + 1 **|** participant +
1 **|** interval, REML = FALSE)

3.5.4 Results

Descriptive statistics for estimates in both active and passive trials, at each interval and according to the voice outcome are given in Table 6. Full model outputs are reproduced in Supplemental Table 12.

Table 6. *Mean interval estimates (ms) across conditions in Experiment 5.*

	Active Trials		**Passive Trials**	
INTERVAL	Self-voice	Other-voice	Self-voice	Other-voice
100ms	152	174	152	187
400ms	317	350	347	382
700ms	514	549	559	589

The interval estimates are plotted in Figures 13a and 13b. The interaction between *agency* and *voice identity* was non-significant, as in Experiment 4. However, there was again a significant main effect of *agency* ($\chi 2(1) = 23.439$, $p = <.001$) showing that interval estimates were significantly lower in the active trials than the passive trials ($p<.001$) across both self- and other-voice outcomes. Thus, in trials in which participants made an action to generate an outcome, they perceived the interval between that action and its outcome to be temporally compressed relative to the same interval when no starting action was made. This result, in line with previous temporal binding studies, suggests participants had a greater sense of agency in active trials.

Further, there was a significant main effect of *voice identity* ($\chi 2(1) = 37.985$, $p = <.001$) showing that intervals were perceived as being shorter in duration when they terminated in the self-voice as compared to the other-voice ($p=<.001$). Here again, the effect of *voice identity* modulated interval estimates in both active and passive trials. Thus, the relevance of an outcome to the self appears to modulate temporal perception even in conditions that cannot be underpinned by a sense of agency. This, again, calls into question whether the self-associated nature of an outcome can be said to influence agency, per se, or whether it affects temporal judgements more broadly which are, here, taken as an implicit measure of agency.

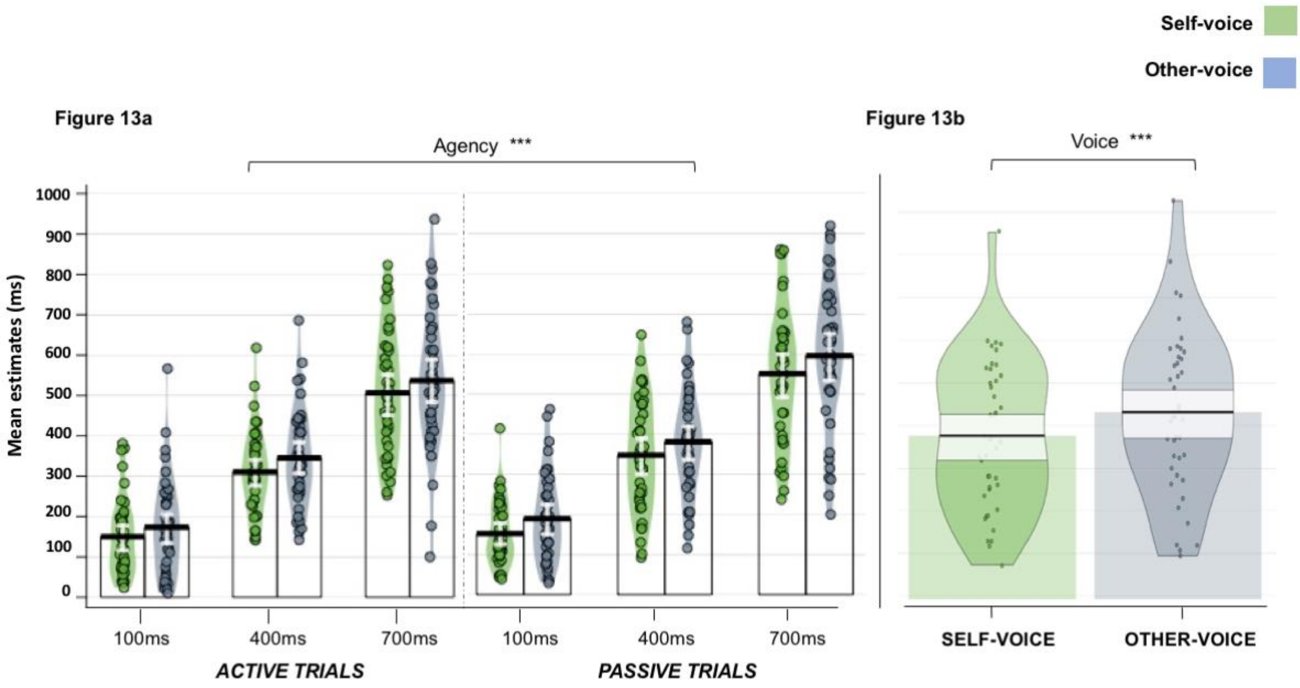

Figure 13a. Experiment 5**:** Mean interval estimates (ms) according to the outcome (self-voice or other-voice), the true interval duration (100ms, 400ms, and 700ms), in both active and passive trials. The error bars indicate the SEs of the means. Coloured segments show smooth density curves for the full data distribution, while individual dots indicate mean estimates per participant. **Figure 13b.** Experiment 5**:** Mean interval estimates (ms) according to the outcome (self-voice or other-voice), collapsed across all conditions. Inference bands indicate CIs. Horizontal bars show main effects with asterisks denoting significance within estimates as determined by likelihood ratio tests. *** p<.001

3.6 General Discussion

In Experiments 4 and 5, I used an interval estimation task to measure temporal binding as an implicit measure of sense of agency. I hypothesised that there would be stronger temporal binding between actions and outcomes if the outcome was a self-owned voice relative to an other-owned voice. Further, that this increased binding for the self-voice – and so agency over it – could be used to corroborate whether that voice had been sufficiently integrated into the self-concept so as to be causally linked to self-action. The results across these two experiments are mixed.

First, it is apparent that the pattern of results, in which interval estimates are lower for intervals preceding a self-voice outcome than an other-voice outcome – is evident in both Experiments 4 and 5. Interestingly, the magnitude of the effect of self-ownership increased from Experiments 4 to 5. This is reflected in the greater significance level of the effect in Experiment 5 (Experiment 4: $p=.008$; Experiment 5: $p<.001$), the higher model estimates (Experiment 4: $\chi2(1) = 6.837$; Experiment 5: $\chi2(1) = 37.986$), and in the average difference in estimates in milliseconds for intervals terminating in the self-voice relative to the other-voice (Experiment 4: 14ms; Experiment 5: 41ms). This may be due to the additional postdictive identity judgement in Experiment 5 which emphasised whether an outcome was self-relevant or not. Indeed, the only difference between the two experiments – other than the sample – was the addition of a postdictive judgement task in which participants reported the ownership status of the voice. It is possible then that by asking participants to state who the voice belonged to, the prominence and self-relevance of the self-owned voice was increased relative to the other-owned voice. In turn, the magnitude of the effect of voice identity (i.e. whether it was self-owned or not) on temporal binding increased.

This demonstrates that by manipulating participants' level of engagement with the cues signalling whether a voice was self-owned, or other-owned, the identity of the outcome postdictively modulated participants' temporal perception. However, how exactly self-association affects interval estimates remains unclear, given that it influenced both active and passive trials. Currently, there is only one other study that has analysed temporal binding with explicitly self-associated outcomes in comparison to outcomes associated with others. Makwana & Srinivasan (2019) ran an interval estimation task in which participants made a keypress and then judged the duration of an interval before visual outcomes (i.e., geometric shapes) that had previously been associated with self, friend, or other. The authors found shorter estimates for intervals preceding the self-associated shape relative to the other identities and thus concluded that self-associated stimuli produce stronger temporal binding (Makwana & Srinivasan, 2019). However, the authors did not use a passive condition and so the results of the current study may be the first to demonstrate that self-association also affects interval estimates in conditions of no agency. Moreover, that the relatively lower interval estimations for self-associated outcomes may not be indicative of a greater sense of agency over those outcomes as this result occurs under a condition of no agency.

Notably, several previous studies have investigated the factors which may modulate sense of agency by assessing differences in interval estimates according to these factors within active trials only (Barlas & Kopp 2018; Obhi, Swiderski, & Brubacher, 2012; Ulloa, Vastano, George, & Brass, 2019; Zopf, Polito, & Moore, 2017.) Within such studies, the conditions under which interval estimates are lower are deemed to show greater sense of agency. However, Experiment 4 and 5 demonstrate that it is also possible that such factors can also affect the interval estimates in the passive trials. Given that the passive trials cannot be underpinned by agency, it is possible that these factors are not interacting with sense of agency but with temporal judgements more widely. This finding may therefore serve as the basis for future research which must work to elucidate what the presence of temporal binding in passive trials may reflect.

If the influence of self-associated outcomes on temporal judgements is not underpinned by a greater sense of agency for the outcome, it remains unclear how they interact. The results from Chapter 2 clearly demonstrate that a new self-voice is perceptually prioritised, being recognised more quickly and accurately in speeded judgement tasks relative to the voices of others. Here then, it could be that the self-voice outcome within the interval estimation task is perceived and processed more quickly because it perceptually prioritised. This quicker recognition of a self-voice once it is heard may, in turn, postdictively modulate the perception of the interval duration preceding it. The pilot study presented in this chapter has already demonstrated that the duration of the cue-tone before the interval can modulate the perception of the interval's duration. Thus, it is possible that the same interval can be postdictively modulated too by the *perceived* duration of the outcome.

Clearly, the finding that participants perceived intervals preceding a self-voice as perceptually shorter than intervals preceding an other-voice does not unequivocally support a greater sense of agency for the self-voice. In spite of this though, these results do show that participants' temporal perception was modulated according to whether the voice was self-owned or other-owned. Given that, prior to this task, both voices were equally unfamiliar, this does corroborate that the voice outcomes differentially affected temporal judgements by virtue of ownership. Thus, the results from this paradigm support and build upon the previous findings presented in Chapter 2: by associating a voice to the self through ownership, people not only prioritise it in perception, but the status of ownership influences temporal perception. Further studies may elucidate the extent to which this biasing of temporal perception is underpinned by a greater sense of agency over that self-voice relative to voices owned by others.

4 Processing the true self-voice versus a new self-voice

Abstract

In Chapters 2 and 3, I explored the way in which an unfamiliar voice can become self-relevant and, thereafter, receive a processing advantage. The fact that this voice has, thus far, been unfamiliar has been necessary to determine the precise benefit conferred to a stimulus because it has become self-relevant, separately to any benefit that might be conferred to a stimulus because it is familiar. Therefore, the extent to which people perceptually prioritise their own, highly familiar, voice has not been examined. Nor has the extent to which people experience a sense of agency over it relative to a new self-voice they have ownership of. These questions form the basis of Experiment 6.

Results from a perceptual matching task show enhanced perceptual prioritisation over the true self-voice relative to a new self-voice that participants have ownership of. Conversely, participants' sense of agency did not increase over their true self-voice relative to a voice they newly owned, as measured in an interval estimation task. These results provide a benchmark against which we can derive how far an other-generated voice can be processed as "self" in comparison to the true self-voice. In Experiment 7, participants were given ownership of a new self-voice while their true self-voice was presented in-task as belonging to an 'other'. Results demonstrated that the enhanced bias for the true-self voice was diminished when it was owned by an 'other', although participants still attributed bias to a new self-voice which was prioritised over others. An interval estimation task also measured participants' sense of agency and revealed that participants experienced greater agency over a voice they owned within the task relative to the true self-voice when it was framed as being 'other-owned'. Together, Experiments 6 and 7 elucidate the extent to which self-bias is flexibly applied according to the context and, further, how participants' sense of agency is influenced dynamically by what they believe is most self-relevant.

4.1 Introduction

The true self-voice differs from a newly self-associated voice in several key ways, all of which may affect the bias afforded to it and the agency experienced over it.

First, the true self-voice is much more familiar and this familiarity may confer its own benefit. For instance, both identity perception (Kanber, Lavan, & McGettigan, 2021), and intelligibility (Holmes, To, & Johnsrude, 2021) are facilitated by familiarity with a voice. This is particularly evident in adverse conditions such as when the acoustic signal is degraded (Xu et al., 2013) or when in background noise (Johnsrude, Mackay, Hakyemez, et al, 2013). Further, familiarity with a voice can also aid the ability to ignore it, such as when trying to listen to one voice while suppressing another in a busy social setting (Johnsrude et al., 2013).

Greater familiarity with the true self-voice may therefore lead to quicker recognition of it relative to a newly self-associated voice. Given that the perceptual matching paradigm determines self-bias according to how quickly a voice is recognised as self-associated or not, this familiarity may interact with the measure of bias; the true self-voice may appear relatively more prioritised than a newly self-associated voice according to faster reaction times and increased recognition accuracy. However, whether familiarity actually increases self-bias, or simply aids voice recognition, will be difficult to disentangle. Thus, it may be that there is greater prioritisation for the true self-voice because its familiarity facilitates recognition but, importantly, that may not reflect greater self-bias for the true self-voice as a self-relevant stimulus.

In the wider literature on self-stimuli, a wealth of studies has tried to resolve the issue of disentangling self-bias and familiarity. For instance, Bortolon & Raffard (2018) ran a meta-analysis of 54 studies focusing only on self-face processing that questioned whether the self-face is processed with self-bias or, rather, just as a highly familiar stimulus. The results demonstrate that the self-face is shown to consistently receive a processing advantage relative to either familiar faces (e.g. faces of family, friends, celebrities) or unfamiliar faces in tasks spanning across attention, perception, and memory. Thus, the processing of the self-face is regarded as distinct from the way we process the faces of other people; there is faster and more accurate processing for the self. This is likely because the greater exposure to the self-face enhances our stored representation of it (Bortolon & Raffard, 2018) and also, due to its strong association with the self, we both prioritise it and regard it more positively (Ma & Han, 2010). Thus, this preferential processing arises both because of the greater familiarity with the stimulus and because of the advantage for self-associated stimuli.

By examining the true self-voice in the perceptual matching paradigm then, I purposefully confound familiarity and self-bias. The aim here is not to disentangle them but, rather, to use the paradigm to derive a degree of perceptual prioritsation for the true self-voice relative to a newly self-associated voice. If there is enhanced prioritisation for the true self-voice, it will remain unclear whether this is driven by greater self-bias or by familiarity. However, it is still important to explore because when the true self-voice is perceived in reality, it is never stripped of its familiarity. This study can therefore provide insight into how the true self-voice is perceived relative to others and provide a benchmark of prioritisation against which the processing of a newly self-associated voice may be compared.

Of course, it should also be noted that perception of the true self-voice within this study relies on recognising it on a recording. It is well known that passively hearing the self-voice on a recording is acoustically different from how it sounds when we speak aloud. When we speak aloud there is additional vibrational input via bone and skin conductance (Békésy, 1954; Maurer & Landis, 1990) which typically means the voice sounds deeper and richer (Kimura & Yotsumoto, 2018) compared to when the voice is played back through headphones or a loudspeaker. Despite this perceptual disparity, studies have shown that people are good at recognising their self-voice on a recording, with recognition accuracy at 89-93% in a study by Hughes and Nicholson (2010) and 94-96% in a study by Rosa et al. (2008). Therefore, although the pre-recorded self-voice may be less familiar to participants than the voice they hear when actively speaking, it remains more familiar than a newly self-associated voice.

The familiarity of the true self-voice may also affect participants' sense of agency over it. In comparison to a newly self-associated voice, participants have extensive prior experience of having agency and flexible control over the true self-voice. In the previous chapter I showed that participants perceived an interval between making an action (i.e., a keypress) and hearing the self-voice to be significantly shorter than an interval between making the action and hearing another's voice. This is, typically, indicative of temporal binding and interpreted as there being a greater sense of agency for the self-voice relative to the other. The true self-voice is one that participants have extensive experience of self-producing. Therefore, this wealth of experience participants have of using their true self-voice – which requires agency, motor control, and intentionality – may be reflected in the temporal binding paradigm as a greater sense of agency over that voice relative to a newly self-associated voice.

Relatedly then, both bias and agency over the true self-voice may differ precisely because the true self-voice is a signal that participants have had agency in using and one that has been previously self-generated. Studies have shown that self-generated outcomes are encoded differently; we recognize writing we ourselves have written more easily than if someone else wrote it (Knoblich, Seigerschmidt, Flach, & Prinz, 2002). Also, we remember words we have said aloud better than words we have only read (Maslowski, Meyer & Bosker, 2018). It is possible therefore that simply by being self-generated, participants are more biased towards the true self-voice, and experience greater agency over it because it is an outcome they have previously self-produced.

Finally, the true self-voice may be more self-representative and identity-congruent. In being the inherent biological property of an individual, the true self-voice conveys a multitude of unique identity cues. Previous work examining how the identity relevance of a to-be-associated stimulus can modulate how prioritisation of it, and agency over it, is limited. Experiment 2 within this thesis asked whether gender-matching a new voice to the self would make it more identity-relevant and so increase perceptual bias towards it. The results of that experiment suggested that the identity-relevance of to-be-associated stimuli did not modulate prioritisation. As such, the results suggested that only the concept to which a new stimulus was associated (i.e., self, friend, other) modulated the degree of prioritisation with which the stimulus was processed. This mirrors the finding of Golubickis et al., (2019) who demonstrated that prioritisation can be modulated by the identity relevance of group concepts to which the stimulus is associated (i.e., musician, athlete, vegan). In this instance, although all three concepts were identity-relevant, the stimulus that was associated with the concept of the highest identity relevance was prioritised.

All of these factors – a voice's self-generation, its greater familiarity, its maximal identity-relevance – may make the true self-voice more self-relevant than a newly self-associated voice that is only self-relevant through a participant's ownership of it. Therefore, assessing how people experience their true self-voice could provide a critical benchmark against which we can derive the extent to which an unfamiliar stimulus has become self-associated. This benchmark can be used to corroborate our previous results. For instance, from the studies presented in Chapter 2 using the perceptual matching paradigm, I conclude that unfamiliar, external items such as a new voice have become self-associated. This conclusion is based solely on faster reaction times and improved accuracy for the self-associated stimulus relative to a less self-relevant stimulus (i.e., a voice belonging to a friend or an other). Thus far, it could be concluded that any self-owned stimulus will be prioritised equally regardless of: a) its level of familiarisation, b) how representative of self it is and, therefore c) how self-relevant it is. By investigating self-bias over the true self-voice, we can provide key insights into the maximum prioritisation a stimulus may receive.

Similarly, by examining how participants' sense of agency may differ between the true self-voice and a self-associated voice, it will allow the previous temporal binding results to be better placed within this context. The true self-voice is a signal that can only exist if a participant has previously made an intentional action – requiring agency – to use it. Thus, it may be that this is reflected in participants sense of agency over the true self-voice outcome in comparison to a voice they have not had agency in producing.

4.2 Experiment 6

Here I investigated self-bias for the true self-voice within two tasks. First, I assessed the extent to which participants perceptually prioritised their true self-voice relative to a new self-owned voice, using a perceptual matching paradigm. Second, I assessed the sense of agency participants experienced over it relative a self-owned voice, using a temporal binding paradigm.

4.2.1 Participants

96 participants (mean age = 26.8 years, SD = 6.11 years, age range = 18–40, 48 female, 48 male) completed the study. This data set of 96 arose following the exclusion and replacement of 7 participants based on exclusion criteria outlined below. Participants were recruited online via Prolific (www.prolific.ac) as British monolingual speakers of English with no visual impairments or hearing difficulties. Only participants with over 90% approval rate on the recruitment platform were eligible. Participants were also required to use Google Chrome as their internet browser and to pass a headphone check (Milne et al., 2020) to ensure they were wearing headphones. Ethical approval was obtained and all participants gave informed consent prior to testing.

Of these 96 participants, 48 participants (mean age = 27.2 years, SD = 6.35 years, age range = 19-40, 24 female, 24 male) were invited to an initial voice recording session run remotely online via Gorilla.sc (see Stimuli for full details). Experimental participants are therefore those which heard their true self-voice within the tasks. The remaining 48 participants acted as controls; all three voices heard within the tasks were unfamiliar to them., Each control participant was gender-matched to an experimental participant who had previously recorded their voice.

4.2.2 Methods

4.2.2.1 Stimuli

Self-voice stimuli

Participants were asked to ensure they were in a quiet environment with minimal background noise and were required to use a laptop/desktop PC with a working microphone to record their voice.

Within this voice recording task, participants were first asked to test their microphone by recording a short clip of their voice and playing it back aloud in Gorilla.sc.

Participants had to report: a) whether they could hear the recording; b) whether it was clear; c) whether it was loud enough, before continuing to the main task.

The purpose of this task was to obtain high quality recordings of each participant saying "hello" as naturally as possible. The target word, "hello", was therefore embedded in various sentential contexts (see Appendix), which participants were asked to read aloud across 12 different trials. Within each trial, the text appeared and remained on screen for 20 seconds, ensuring participants had time to comprehend and read aloud the required sentence. Alongside the appearance of the text, participants heard a voice counting down (e.g., "3…2…1…") and participants were asked to start reading the text at the end of the countdown. This ensured time for the recording channel to open and begin recording before the onset of the speech.

The purpose of the first 9 trials was to allow participants time to become accustomed to recording their voice, such that their voice might sound more natural by trials 10-12. In these final trials, participants were asked to repeat the word "hello" ten times per trial, ensuring they took a breath between each iteration. First, participants were asked to repeat "hello" in a cheerful voice, then in a neutral voice, and finally a sad voice.

Participants were then debriefed and asked to report any problems or difficulties they experienced in completing the task. All participants were paid for their participation and informed that they would be contacted if eligible to complete the main study. Specifically, if all instructions had been followed and if their recordings were free of background noise and of sufficient quality for use. Of these 48 participants, only 3 produced voice recordings that were unusable. These participants were paid for their participation but replaced within the study with a further 3 participants.

Friend- and other-voice stimuli

A further 4 participants (mean age = 31.8 SD = 2.2 years, range = 29-35, 2 female) were recruited to record their voice. Recordings of "hello" from these four voices were used within the main experiment's tasks as either the 'friend' or 'other' voice. These participants were not invited to participate in any subsequent sessions.

Pre-processing voice stimuli

The target word "hello" was extracted from each participant's recordings, including from the recordings collected to be used as the friend-voice and other-voice. In each case, the voice clip was extracted from the trial in which participants repeated the target word in a neutral voice. This was done to ensure as much uniformity as possible across participants in terms of the affective valence of the voice. All stimuli were normalized for RMS amplitude using PRAAT (Boersma, 2001; Boersma and Weenink, 2020).

4.2.2.2 Procedure

Perceptual Matching Task

First, participants completed a perceptual matching task (for full procedure, see Chapter 2, section 2.2.3.2) to examine perceptual prioritisation over the self-voice. Briefly, this task comprised a familiarisation phase and a test phase. In the familiarisation stage all the participants were told that within the task they would hear three different voices. For the experimental participants this was: "*Your own voice (recorded from Session 1) which belongs to you, a voice which belongs to a friend, and a voice which belongs to a stranger".* For the control participants this was: *A voice which belongs to you, a voice which belongs to a friend, and a voice which belongs to a stranger".*

Each experimental participant heard their own pre-recorded voice token as the self-voice, labelled as "You". All test voices were gender matched to the participant (e.g. if the self-voice was female, the two voices heard as friend and other (i.e. stranger) were also female), and which of these two voices was assigned to each identity was counterbalanced across participants. Each control participant was then randomly matched to an experimental participant according to their gender. That control participant heard the same three voices associated to each self, friend, and other as the corresponding experimental participant. This ensured that the only difference between the two participants' experience was whether the self-voice was, truly, the participant's true self-voice that had been pre-recorded or whether it was a newly self-associated voice. Within the test phase, the visual labels denoting the voice identities were the same across all participants (both experimental and control): either 'YOU', 'FRIEND' or 'STRANGER'.

Participants performed three blocks of 72 trials (216 trials in total) and were informed of their overall accuracy at the end of each block.

Interval Estimation Task

Participants then completed an interval estimation task to measure participants' sense of agency over the self-voice and the other-voice (for full procedure, see Chapter 3, section 3.2.1).

In the active condition, the first screen displayed "Waiting…" and the participant was instructed to press ENTER at any time of their choosing. When they pressed ENTER, there was an interval lasting either 100ms, 400ms, or 700ms. After the interval, one of two possible auditory stimuli played; either a clip of the self-voice or of the voice they had been told belonged to an "other". Similarly, in the passive condition, the "Waiting…" screen appeared until a 100ms tone passively marked the onset of the interval, lasting either 100ms, 400ms, or 700ms as in the active trials. After the interval, participants heard either the self-voice or the other-voice. In both conditions, participants had to respond as to whether the voice-outcome they'd heard was the voice that belonged to them, or the voice that belonged to the other. Critically then, for the experimental participants that self-voice was the true self-voice, and for the control participant it was a newly self-associated voice.

There were four blocks: two blocks of active trials and two blocks of passive trials. The order of the blocks was randomised across participants. In each block, there were 30 trials (120 total) such that each voice (self, other) was presented 10 times after each interval duration (100ms, 400ms and 700ms, with randomised presentation), in both active and passive conditions. After the interval estimation task, participants were debriefed and paid for the participation. The whole study took participants 35 minutes to complete, on average.

4.2.2.3 Pre-processing

Perceptual Matching Task

I analysed three measures: reaction times, sensitivity, and accuracy. Reaction times were measured as the delay between the onset of the visual label and the participant's response as categorised by the *voice identity* (self-voice, friend-voice, other-voice) per trial. To measure sensitivity, unbiased *d'* scores were calculated by combining performance scores from match and mismatch trials at each level of the *voice identity* factor (self, friend, other). Specifically, YES/NO responses were recoded into hits and false alarms, following the log- linear approach to adjust for cases involving 100% hits or 0% false alarms (Stanislaw & Todorov, 1999).

Interval Estimation Task

For each participant, a mean interval estimation per participant was calculated in each condition e.g., Active trials with a 100ms interval were presented 10 times per participant and a mean estimate was calculated across these 10 trials. This mean estimation was used only to apply the exclusion criteria (see below).

4.2.2.4 Exclusion Criteria

Perceptual Matching Task

All erroneous responses, as well as responses shorter than 200ms or longer than 1500ms were removed from the RT analysis in line with Sui et al.'s (2012) approach. This accounted for only 1.1% of trials. Whole datasets for 3 (1 experimental, 2 control). participants showing overall performance accuracy at chance (≤50% + 95% CI) were excluded as this indicated random responses and/or an inability to distinguish between the self-voice and the other's voice.

Interval Estimation Task

In the interval estimation task, whole datasets of participants whose mean estimates did not increase monotonically with the presented interval duration (100ms, 400ms, 700ms) were also excluded and replaced as this demonstrates particularly difficulty with the task and/or a lack of engagement (Suzuki et al., 2019; Caspar et al., 2016). A further 4 (1 experimental, 3 controls) participants were excluded on these criteria by checking their average interval estimation increased numerically with the interval duration.

Across the perceptual matching task and the interval estimation task then, a total of 7 participants were excluded and their data was replaced until a total of 96 participants had successfully completed the study. Lastly, I eliminated all trials more than 2 standard deviations above or below each participant's mean (per condition) in line with previous studies (Engbert et al., 2008) as this demonstrated erratic performance. <1% of trials were eliminated on these criteria.

4.2.2.5 Analysis

Perceptual Matching Task

The perceptual matching task consisted of one within-subject factors of 'voice identity' (*self* vs *friend* vs *other*) and one between-subjects factor of 'condition' (experimental: *true self-voice* vs control: *new self*-voice). The within-subjects factor of 'trial type' (*match* vs *mismatch*) was also present.

Both reaction time data and sensitivity were assessed with linear mixed models (LMM) using lme4 (Bates et al., 2014) in the R environment (R Core Team, 2013). In either analysis, I ran a model that included an interaction between *condition* and *voice identity* and a random intercept of 'participant'. Reaction time data was analysed in match trials and, separately, mismatch trials.

E.g. lmer(reaction time ~ condition * voice identity + 1 | participant, REML = 'FALSE')

To assess accuracy, I ran a binomial generalized linear mixed model (GLMM) that, again, included an interaction between *condition* and *voice identity* and a random intercept of *participant,* in match trials and, separately, mismatch trials.

E.g. glmer(accuracy ~ condition * voice identity + 1 | participant, family = "binomial")

Interval Estimation Task

The interval estimation task consisted of three within-subject factors: 'voice identity' (*self* vs *other*), 'agency' (*active* vs *passive*) and 'interval' (*100ms* vs *400ms* vs *700ms*), as well as one between-subjects factor of 'condition' (*true self-voice* vs *new self*-voice).

If there is a greater sense of agency over a true self-voice relative to a new self-voice, this would be indicated by an interaction between *agency, condition,* and *voice identity.* Specifically, estimates for the intervals terminating in the true self-voice would be lower than estimates for intervals terminating in the new self-voice in active trials only, while interval estimates for the other-voice should not differ across *condition*. If participants have a greater sense of agency over a self-associated voice, but this agency is unaffected but whether that self-associated voice is the true self-voice or not, this would be indicated by an interaction between *voice identity* and *agency.*

Interval estimations were assessed with linear mixed models that first included an interaction between *agency, condition,* and *voice identity.* In all analyses, statistical significance of the effects was established via likelihood ratio tests by dropping effects of interest from the appropriate model. For example, to establish whether an interaction was significant, the interaction was dropped from a model including the fixed effects. To test for the significance of the main effects, the relevant effect was dropped from a model that only included the other main effects. For all analyses reported here, post-hoc comparisons were conducted in emmeans (Lenth, 2016). All post-hoc comparisons of main effects were adjusted via Bonferroni correction.

E.g. lmer(interval estimation ~ agency * condition * voice identity + 1 | participant + 1 | interval, REML = FALSE).

For GLMMs, odds ratios and associated confidence intervals are provided as an estimate of the size of the relevant effects. An odds ratio of 1 indicates that no effect is present, while odds ratios that deviate from 1 indicate an effect is present. The larger the deviation, the bigger the effect. If the confidence intervals do not cross 1, the relevant effect is significant. For LMMs, the models' estimates and associated confidence intervals for each effect are reported. The further away from 0 the estimates are, the bigger the effect. If the confidence intervals do not cross 0, the relevant effect is significant.

4.2.3 Results

4.2.3.1 Perceptual Matching Task

Descriptive statistics for reaction times, accuracy, and sensitivity scores (*d'*) across conditions are given in Table 7.

Table 7. *Mean RTs, accuracy, and sensitivity (d') in Experiment 6 (match trials)*

Condition	Voice Identity	Mean RT (ms)	Accuracy	*d'* *
True self-voice	Self	459 (148)	0.98 (0.12)	4.07 (0.49)
	Friend	529 (201)	0.97 (0.18)	3.73 (0.68)
	Other	571 (222)	0.92 (0.26)	3.37 (0.99)
New self-voice	Self	507 (190)	0.94 (0.24)	3.23 (0.89)
	Friend	549 (207)	0.93 (0.25)	3.11 (1.01)
	Other	571 (214)	0.90 (0.30)	2.86 (1.01)

Note. RT = reaction time; Accuracy - proportion correct. Standard deviations appear within parenthesis. *Performance scores from match and mismatch trials are combined to provide d' scores.

Reaction times

The reaction time data from the matched trials are plotted in Figure 14. There was a significant interaction ($\chi2$ (2) = 31.13, p= <.001) between *condition* (*true self-voice* vs *new self*-voice) and *voice identity* (self, friend, other). Participants' reaction times for the self-voice were significantly faster when that voice was the true self-voice relative to an assigned self-voice (p<.001). By comparison, the reaction times for the friend-voice did not differ between groups (p=.09), nor did reaction times for the other-voice (p=.11). This demonstrates that the self-voice was prioritised more highly if it was the true self-voice than if it was a *new* self-voice that had been randomly assigned to them. The post-hoc comparisons also revealed a significant self-prioritisation effect in both conditions; there were significantly quicker reaction times to the self-voice relative to either the friend-voice (p<.001) or the other-voice (p<.001). Further, reaction times were quicker to the friend-voice relative to the other-voice in both conditions (p<.001).

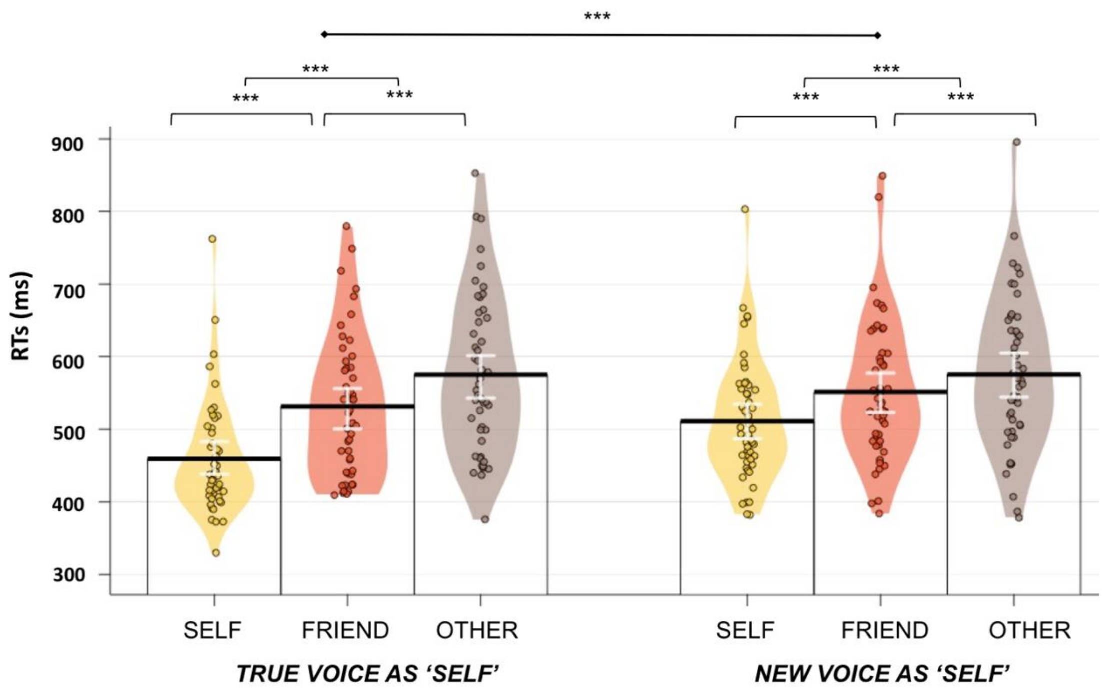

Figure 14. Experiment 6**:** Mean reaction times (RT) as a function of the voice identity (self, friend, or other) and voice type (true self-voice vs. new self-voice). **Graph models match trials only.** The error bars indicate the SEs of the means. Coloured segments show smooth density curves for the full data distribution, while individual dots indicate mean RTs per participant. The top bar indicates the significance of the interaction, horizontal bars show post-hoc comparisons. Asterisks denote significance within RTs as determined by likelihood ratio tests. *** $p<.001$

These results showed that, regardless of whether the self-voice was the true self-voice or not, participants prioritised it relative to the others. This is in line with the previous findings but I show here, for the first time, enhanced prioritisation for the true self-voice relative to a newly assigned self-voice. Full model outputs for this analysis are reproduced in Supplemental Table 13.

The data from the mismatch trials are plotted in Figure 15. There was a significant interaction between *condition* and *voice identity* ($\chi2(2) = 14.463$, $p<.001$). Again, experimental participants – for whom the self-voice was the true self-voice – responded more quickly to the self-voice relative to control participants for whom the self-voice was newly assigned ($p<.001$). However, both experimental participants and control participants had similar reaction times for both the friend-voice ($p=.678$) and the other-voice ($p=.491$).

Additionally, the post-hoc comparisons showed that RTs for the self-voice were significantly quicker than for either the friend-voice (p<.001) or the other-voice (p<.001) in the experimental group, although RTs to the friend-voice did not significant differ from RTs for the other-voice (p=.711). Finally, in the control group, RTs for the self-voice were quicker than RTs for the friend-voice (p<.001) but not the other-voice (p=.456) and RTs for the other-voice were significantly quicker than for the friend-voice.

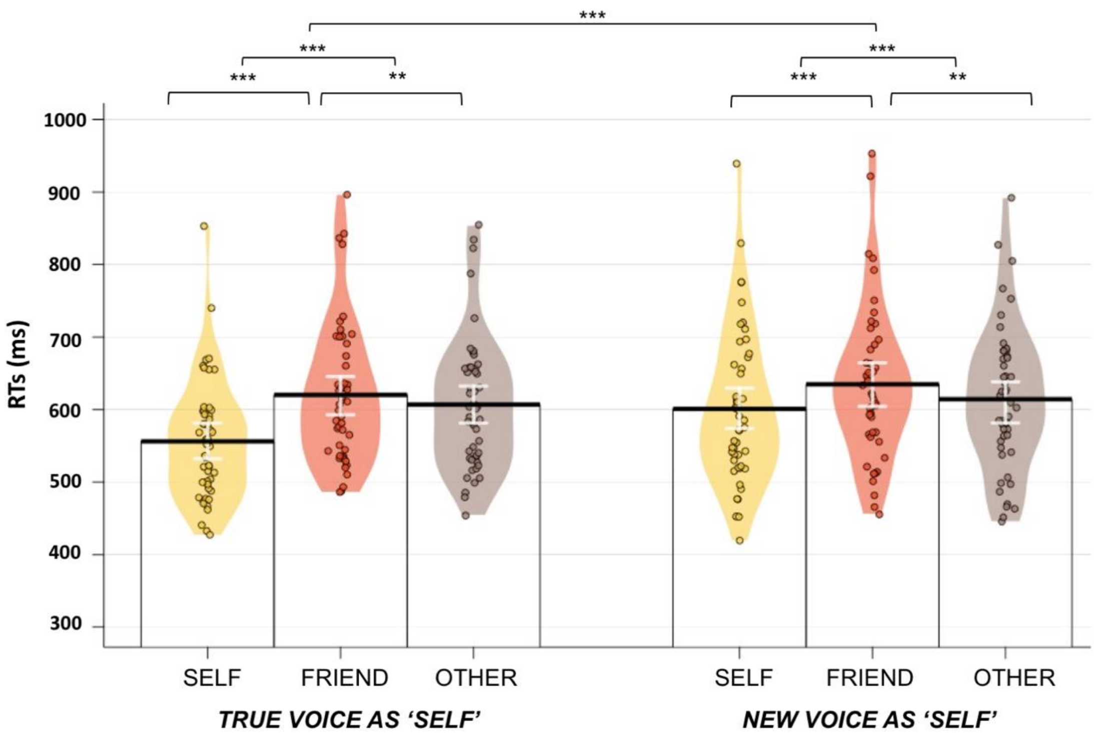

Figure 15. Experiment 6**:** Mean reaction times (RT) as a function of the voice identity (self, friend, or other) and voice type (true self-voice vs. newly-associated self-voice). **Graph models mismatch trials only.** The error bars indicate the SEs of the means. Coloured segments show smooth density curves for the full data distribution, while individual dots indicate mean RTs per participant. The top bar indicates the significance of the interaction, horizontal bars show post-hoc comparisons. Asterisks denote significance within RTs as determined by likelihood ratio tests. *** p<.001

Overall, the RT data suggest that a self-voice is consistently prioritised, but this prioritisation is enhanced if it is the true self-voice compared to a newly self-associated voice.

Sensitivity

Full model outputs for this analysis are reproduced in Supplemental Table 14. The interaction between *condition* and *voice identity* was non-significant ($\chi 2(2)=2.985$, $p=.225$) here but there was, again, a significant main effect of *voice identity* ($\chi 2(2)=27.68$, $p<.001$). Participants showed increased perceptual sensitivity to the self-voice relative to the other-voice ($p<.001$) and, further, for the friend-voice relative to the other-voice ($p=.007$). Thus, the self-voice was prioritised regardless of whether it was the true self-voice or a new voice that had become self-associated. However, perceptual sensitivity did not significantly differ between the self-voice and the friend-voice ($p=.066$) for either group. The main effect of *condition* was also significant ($p<.001$); participants who heard their true self-voice as 'self' were more perceptually sensitive to that self-voice ($p<.001$), but also to the friend-voice ($p<.001$) and the other-voice ($p<.001$) relative to controls who were newly assigned a self-voice. Overall, this shows that a self-prioritisation effect was evident in both groups, with the experimental participants having higher perceptual sensitivity to all voices, not just the self-voice. Taken together with the RT results, this suggests that when the self-voice is the true self-voice, participants are perceptually more sensitive to this voice and to the other voice identities. This may be because participants were better at discriminating what was – and was not – a match with the self, when the true self-voice was more familiar. This may have led to higher accuracy in the task generally, across all three voices.

Accuracy

Last, a GLMM on trialwise accuracy was run. Full model outputs are reproduced in Supplemental Table 15. The interaction between *condition* and *voice identity* was significant ($\chi 2(2)= 22.09$ $p<.001$). Post-hoc comparisons revealed that participants who heard their true self-voice were significantly more accurate in judging that self-voice ($p<.001$) and also the friend-voice ($p<.001$) relative to controls, though they did not differ in their accuracy of judging the other-voice ($p=.01$, which is non-significant with alpha set at .005 for multiple comparisons). Still, both groups showed a self-prioritisation effect in which accuracy was significantly higher for the self-voice relative to either the friend-voice ($p=.003$) or the other-voice ($p<.001$). Further, accuracy for the friend-voice was significantly higher than accuracy for the other-voice ($p<.001$) in both groups.

The mismatch trials mirror the match trials; the interaction between *condition* and *voice* was again significant ($\chi2(2)= 15.739$ $p<.001$). Post-hoc comparisons revealed that participants who heard their true self-voice were significantly more accurate in judging the self-voice ($p<.001$), the friend-voice ($p<.001$) and the other-voice ($p<.001$) relative to controls. Further, only the experimental participants showed a self-prioritisation effect; only when the self-voice was the true self-voice were participants significantly more accurate at judging that self-voice relative to the friend-voice ($p=.001$) and the other-voice ($p<.001$). The control participants by comparison did not show a self-prioritisation effect in the mismatch trials; accuracy did not significantly differ between the self-, friend-, and other-voice ($p>.05$) when the self-voice was newly associated.

Overall, there was also a self-prioritisation effect across all three measures, showing that a self-associated stimulus will be prioritised perceptually as self-relevant. Critically though, this prioritisation was enhanced in experimental participants for whom the self-voice was their true self-voice. Interestingly, the presence of the true self-voice within the task also increased accuracy and perceptual sensitivity to all three of the voices. This may be both because participants were better able to recognise what was and was not the self-voice amongst the other identities, but also because these participants may have been more engaged in the task with a voice as salient as the true self-voice being presented.

4.2.3.2 Interval Estimation Task

I then assessed how interval estimates were affected by *agency, condition,* and *voice identity* (Figures 16a and 16b). Descriptive statistics for estimates across conditions are given in Table 8.

Table 8. *Mean interval estimates (ms) across conditions in Experiment 6.*

	Active Trials				**Passive Trials**			
	True self-voice		New self-voice		True self-voice		New self-voice	
INTERVAL	**Self**	**Other**	**Self**	**Other**	**Self**	**Other**	**Self**	**Other**
100ms	136	156	154	158	187	200	241	230
400ms	293	312	312	334	385	383	393	416
700ms	514	527	521	552	610	593	609	614

The three-way interaction between *agency, condition*, and *voice identity* on interval estimates was non-significant. Likewise, the two-way interaction between *condition* and *voice identity* was non-significant, which showed that the difference in interval estimates between the self-voice and the other-voice did not differ according to whether the self-voice was the true self-voice or a new self-owned voice. Similarly, the interaction between *agency* and *condition* was non-significant (p=.16), which showed that the difference in interval estimates in active and passive trials did not differ according to condition. Finally, there was a significant interaction between *agency* and *voice identity* ($\chi 2(1)= 6.63$, $p = .01$). Post-hoc comparisons showed that estimates for intervals terminating in the self-voice were significantly lower than for the other-voice but only in the active trials ($p<.001$) and not passive trials ($p=.9$). This shows that the effect of whether a voice was self-owned or not, only affected temporal judgements in the active trials. This is key, because it suggests that self-ownership only affected the interval estimates under conditions in which participants had agency. Finally, post-hoc comparisons showed that there was a typical temporal binding effect in which estimates were significantly lower (for both the self-voice and the other-voice) in active trials relative to the passive trials ($p<.001$). Critically here, these results were unaffected by whether the self-voice was the true self-voice or a newly associated self-voice. Full model outputs are reproduced in Supplemental Table 16.

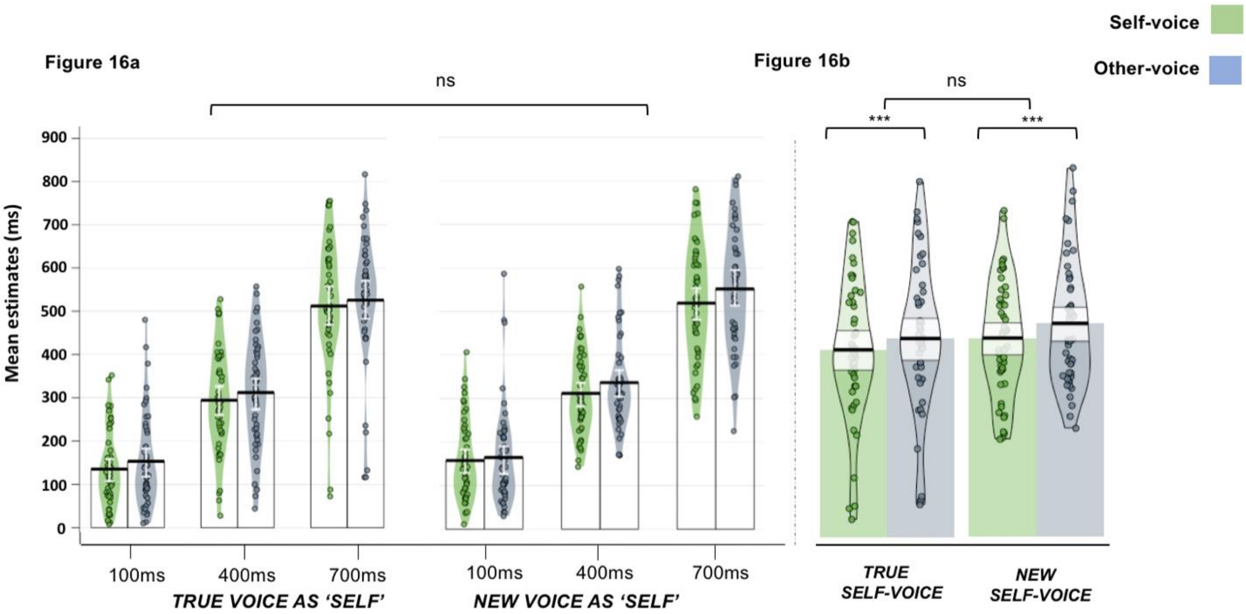

Figure 16a. Experiment 6**:** Mean interval estimates (ms) according to the outcome (self-voice or other-voice), and the condition (true voice as 'self' vs. a new voice as 'self'). Intervals are plotted according to the true interval duration (100ms, 400ms, and 700ms)**.** The error bars indicate the SEs of the means. **Figure 16b.** Mean interval estimates (ms) according to the outcome (self-voice or other-voice) Estimates are collapsed across interval duration. Graphs model active trials only. Inference bands indicate CIs. Coloured segments show smooth density curves for the full data distribution, while individual dots indicate mean estimates per participant. The top bars show interactions, horizontal bars show post-hoc comparisons. Asterisks denote significance within estimates as determined by likelihood ratio tests. *** p<.001

Overall, these results show first that the interval estimates were significantly shorter in the active trials relative to the passive trials. This finding is in line with the wider temporal binding literature and demonstrates a temporal binding effect: people had a greater sense of agency over outcomes they self-generated relative to passively observed outcomes. However, this result comes with the previous caveat that the difference between active and passive trials may also be influenced by differences in task structure across conditions. Second, the results also showed that interval estimates were significantly lower for the self-associated voice outcome than the other-associated outcome, but only in the active trials. Thus, it can be inferred that the effect of *voice identity* did modulate participants' sense of agency rather than temporal perception more broadly. This supports the hypothesis that participants had a greater sense of agency over a self-voice than an other-voice. Finally, and in spite of this sensitivity to whether an outcome was self-relevant or not, agency did not differ according to whether the outcome was the true self-voice or a new self-owned voice. This indicates that, within the task, participants experienced no greater sense of agency over the true self-voice than a newly self-associated voice. These results are discussed in greater detail below.

4.2.4 Discussion

In Experiment 6, I demonstrated a typical self-prioritisation effect over a self-voice, but here showed that it is enhanced if it is the true self-voice relative to a newly self-associated voice. This enhanced prioritisation must be a function of the difference between these two self-voices. Thus, it is possible that the enhanced prioritisation is due to the greater familiarity of the true self-voice, the fact that it is self-generated, and/or that it is more self-relevant.

It may not be possible to disentangle these factors without further study. However, it has already been extensively demonstrated that self-bias exists without, and separately to, what is familiar. Therefore, greater familiarity does not necessarily need to equal greater self-bias for it to still influence prioritisation. Indeed, it is possible that increased familiarity also facilitates perceptual processing similarly to self-bias and may, within this task, also lead to greater prioritisation. It stands to reason that the external sensory stimuli we experience frequently enough for them to become highly familiar are likely to be self-relevant and important for us to prioritise in perception.

Unlike familiarity, which may only facilitate perceptual processing without necessarily increasing self-bias, we may attribute relatively more self-bias to a stimulus that has been self-generated. Within this study only the true self-voice had been self-generated, whereas the self-associated voice had been other-generated and then associated to the self. However, it has now been widely shown that we can attribute self-bias to a range of stimuli that we have not self-generated and still perceptually prioritise them. Indeed, the studies presented within this thesis have extensively demonstrated that we can afford self-bias to a stimulus that is inherently other-generated. Clearly then, self-bias is not *dependent* on a stimulus being self-generated but it could be considered that its being self-generated *enhances* the self-bias attributed to it.

Another way in which the true self-voice and the self-associated voice differ is in their level of "*me*-ness," (i.e., Oyserman, Elmore, & Smith, 2012), or how identity-relevant they are. Within this experiment, the self-voice was substantially more identity-relevant for the participant whose voice it actually was, than for the control participant for whom it was newly self-associated. Indeed, only for the participant for whom it was their true self-voice did it accurately – and uniquely – reflect their age, gender, health, and (recent) affective state (Kreiman & Sidtis, 2011). This difference in identity-relevance may underlie the increased prioritisation of the true self-voice. However, this would be in contrast to those of Experiment 2 (Chapter 2) in which a gender-matched voice, that should be more identity-relevant to the participant in terms of being a more accurate means of representation, did not modulate prioritisation.

Overall, the results from the perceptual matching task show relatively enhanced perceptual prioritisation for the true self-voice relative to a newly self-associated voice. This indicates that our perceptual processing is sensitive to whether a self-voice is the true self-voice or not. However, whether this is because the true self-voice is more familiar, more self-relevant, or self-generated remains unclear. Further studies would be needed to disentangle the factor(s) driving the increased prioritisation here.

The interval estimation task was then used to measure participant's sense of agency over a self-voice – either the true self-voice or a new self-associated voice – relative to an other-voice. The results showed that participants' estimates were equivalent across groups; the difference in temporal binding between self-voice outcomes and other-voice outcomes was not modulated by whether the self-voice was the true self-voice or not. This suggests that people felt relatively more agency over the self-voice than the other-voice, but that this agency was not enhanced for the true self-voice relative to a voice that was only self-owned.

Critically, the difference in estimates for the self-voice outcomes and other-voice outcomes (i.e., the effect of *voice identity)* was significant only in the active trials. This is key because it is in the active trials that participants have agency in generating a voice so it is expected that any factors which modulate agency should influence estimates in the active condition only. However, in previous iterations of this task presented in Chapter 3 (Experiments 4 and 5), the effect of *voice identity* was also present in passive trials in which participants had no agency in generating an outcome. The previous result therefore obscured whether the self-relevance of a voice interacted participants sense of agency or temporal perception more widely. The current result supports the former argument. Indeed, if this study had been run in isolation, the results would be interpreted as indicating a replication of a temporal binding paradigm, showing greater sense of agency over the self-voice. Further, that this enhanced sense of agency is unaffected by whether that self-voice is the true self-voice or a newly owned one. However, in light of the inconsistency in the presence or absence of the effect of *voice identity* in the passive trials across the thesis, further investigation is required before the validity of this interpretation can be determined.

It is surprising then that participants felt similar agency over a true self-voice and a newly owned one, and this is for two reasons. First, because people have extensive experience of producing their own voice and of having agency over it. Indeed, the voice stimulus they heard within the task could not exist without their own prior agentic action. The true self-voice, as a possible outcome, should be highly self-relevant as it inherently signals the agent's identity. However, this is not reflected within the results. Second – and in light of the fact that the presence of the true self-voice did not affect agency – it is surprising that interval estimates were sensitive to the self-relevance of the outcome at all. As in Experiments 4 and 5 (Chapter 3), interval estimates were again significantly shorter for the self-voice outcome relative to an other-voice outcome.

Within the task, participants generated both voice outcomes via a key-press in the active trials and so, in actuality, had equal agency over the two voices. Despite this, participants perceived the time between their actions and the outcomes as relatively more compressed when it was the self-voice outcome. This indicates that interval estimates were affected by whether an outcome was self-relevant or not but, seemingly, were not affected by *how* self-relevant it was; the highly, personally familiar true self-voice did not lead to a greater perceived compression of time than that of a new self-owned voice. It is possible then that an outcome being self-associated is sufficient to accept agency over its occurrence. If this is the case, it could be considered that temporal binding does not function on a scale of ‘how much’ agency is experienced over an outcome but, rather, by whether sufficient agency can be inferred to categorise an outcome as being self-generated, or not.

Experiment 6 thus gives rise to further questions. The first question is whether the enhanced bias and perceptual prioritisation for the true self-voice relative to a newly self-associated one is underpinned by greater self-relevance of the true self-voice. If the true self-voice is prioritised relatively more than a self-owned voice because it is more self-relevant, then this prioritisation should be modulated by altering the relevance of the voice within the task. In Experiment 7 then, I explore prioritisation over the true self-voice when it is instead presented as belonging to an other (a ‘stranger’). Information associated with others is less self-relevant and is typically deprioritised in comparison to information associated with the self. By manipulating the ownership of the voice, it may be perceived less relevant to the self within this task than a newly self-owned voice. If so, despite its being inherently self-relevant, prioritisation of the self-voice may be reduced.

The second question is why sense of agency may be sensitive to an outcome being self-relevant or not but, unlike prioritisation, not to the *degree* to which that outcome is self-relevant. By testing sense of agency over outcomes that include both a newly self-associated voice and the true self-voice, it allows us to assess agency in a within-person design. It may be that only when participants experience two outcomes, both of which are self-associated, that the degree to which those outcomes are differentially self-relevant has an influence on agency. Only in this instance may participants experience more or less agency over one outcome relative to the other. Experiment 7 thus examines the difference in perceptual prioritisation and sense of agency over the true self-voice and the self-associated voice *within* participants.

4.3 Experiment 7

Experiment 7 assessed how participants prioritised a highly salient social stimulus (e.g., the true self-voice) when it is associated with a concept less relevant to the self (i.e., to an other). Relatedly, it examined how the two self-associated voices were prioritised when they were competing: one as a newly self-associated stimulus and the other a self-associated stimulus that is also familiar, self-generated, and identity-relevant. Second, I tested participants' sense of agency over both the true self-voice and a newly self-associated voice by including both voices as possible outcomes to their actions.

4.3.1 Participants

96 participants (mean age = 26.4 years, SD = 6.26 years, age range = 18–40, 48 female, 48 male) were recruited online via Prolific (www.prolific.ac) as British monolingual speakers of English with no visual impairments or hearing difficulties. As before, participants were required to have an approval rate of over 90% to be eligible to take part. This final data set of 96 arose following the exclusion and replacement of 6 participants based on exclusion criteria outlined below.

4.3.2 Methods

4.3.2.1 Stimuli

48 of these participants were invited to complete the same voice-recording task as in Experiment 6. Again, the target word 'hello' was extracted from each participant's voice recordings. Four participants submitted voice recordings that were unsuitable for use (i.e., too much background noise) and these participants were paid for their time but then excluded and replaced within the study.

The voice stimuli used in Experiment 6 as the friend-voice and the other-voice were included again. However, in Experiment 7 these were assigned as either the self-voice or the friend-voice while participant's voice recordings were assigned as the other-voice.

4.3.2.2 Procedure

Perceptual Matching Task

Participants completed a perceptual matching task again comprising a familiarisation phase and a test phase (see Chapter 2 for full procedure). In the familiarisation stage all the participants were told that they would hear three different voices within the task: "*A voice which belongs to you, a voice which belongs to a friend, and a voice which belongs to a stranger".* The experimental participants were also advised that the voice they had previously recorded may not be assigned to the 'self' within the task.

Indeed, for every experimental participant, their pre-recorded voice was assigned to the other-voice (i.e., labelled as belonging to 'Stranger'). All experimental voice stimuli were gender-matched to the participant and to their true self-voice. Each control participant was then randomly matched to an experimental participant according to their gender. The control participant heard the same three voices associated to each self, friend, and other as their corresponding experimental participant.

In the test phase, participants performed three blocks of 72 trials (216 trials in total) and were informed of their overall accuracy at the end of each block.

Interval Estimation Task

Participants completed an interval estimation task to measure participants sense of agency over the self-voice and the other-voice (for full procedure, see Chapter 3, section 3.2.1). In this experiment, however, participants were reminded to complete the task keeping in mind the same voice-identity associations they had learnt in the perceptual matching task (i.e., if they heard the true self-voice, this voice still belonged to the 'other' within the task).

4.3.2.3 Pre-processing

As in Experiment 6, I analysed the same three measures within the perceptual matching paradigm: reaction times, sensitivity, and accuracy. For the interval estimation task, I again calculated a mean interval estimation per participant in each condition e.g., active trials with a 100ms interval were presented 10 times per participant and a mean estimate was calculated across these 10 trials. This mean estimation was used only to apply the exclusion criteria (see below).

4.3.2.4 Exclusion Criteria

Perceptual Matching Task

All erroneous responses, as well as responses shorter than 200ms or longer than 1500ms were removed from the RT analysis in line with Sui et al.'s (2012) approach. This accounted for <1% of trials. Whole datasets for 3 participants showing overall performance accuracy at chance (≤50% + 95% CI) were excluded and their data replaced as this indicated random responses and/or an inability to distinguish between the self-voice and the other's voice.

Interval Estimation Task

In the interval estimation task, whole datasets of participants whose mean estimates did not increase monotonically with the presented interval duration (100ms, 400ms, 700ms) were also excluded and replaced as this demonstrated particularly difficulty with the task and/or a lack of engagement (Suzuki et al., 2019; Caspar et al., 2016). A further 3 participants were excluded on these criteria by checking their average interval estimation increased numerically with the interval duration. Their data was replaced until a total of 96 participants had successfully completed the study.

Lastly, all trials which were more than 2 standard deviations above or below each participant's mean (per condition) were eliminated in line with previous studies (Engbert et al., 2008) as this demonstrated erratic performance. <1% of trials were eliminated on these criteria.

4.3.2.5 Analysis

All measures were again assessed via LMM/GLMMs using the same models specified in Experiment 6. The perceptual matching task consisted of two within-subject factors of 'voice identity' (*self* vs *friend* vs *other*) and 'trial type' (*match* vs *mismatch*) and one between-subjects factor of 'condition' (experimental: *true self-voice as 'stranger'* vs control: *unfamiliar voice assigned as 'stranger'*).

The interval estimation task consisted of three within-subject factors: ‘voice identity’ (*self* vs *other*), ‘agency’ (*active* vs *passive*), and ‘interval’ (*100ms* vs *400ms* vs *700ms*) and one between-subjects factor: ‘condition’ (experimental: *true self-voice as ‘stranger’* vs control: *unfamiliar voice assigned as ‘stranger’*).

4.3.3 Results

4.3.3.1 Perceptual Matching Task

Descriptive statistics for reaction times, accuracy, and sensitivity scores (*d’*) for experimental and control participants according to Voice are given in Table 9.

Table 9. *Mean RTs, accuracy, and sensitivity (d’) in Experiment 7 (match trials)*

Condition	**Voice Identity**	**Mean RT (ms)**	**Accuracy**	***d'* ***
Experimental	Self	527 (227)	0.94 (0.24)	3.35 (0.91)
	Friend	548 (218)	0.96 (0.20)	3.60 (0.77)
	Other	541 (209)	0.92 (0.27)	2.98 (0.79)
Control	Self	513 (205)	0.96 (0.20)	3.41 (0.71)
	Friend	546 (211)	0.96 (0.19)	3.45 (0.76)
	Other	551 (222)	0.92 (0.27)	2.81 (1.02)

Note. RT = reaction time; Accuracy - proportion correct. Standard deviations appear within parenthesis. *Performance scores from match and mismatch trials are combined to provide d’ scores.

Reaction times

The RT data from matched trials are plotted in Figure 17 and full model outputs are reproduced in Supplemental Table 17.

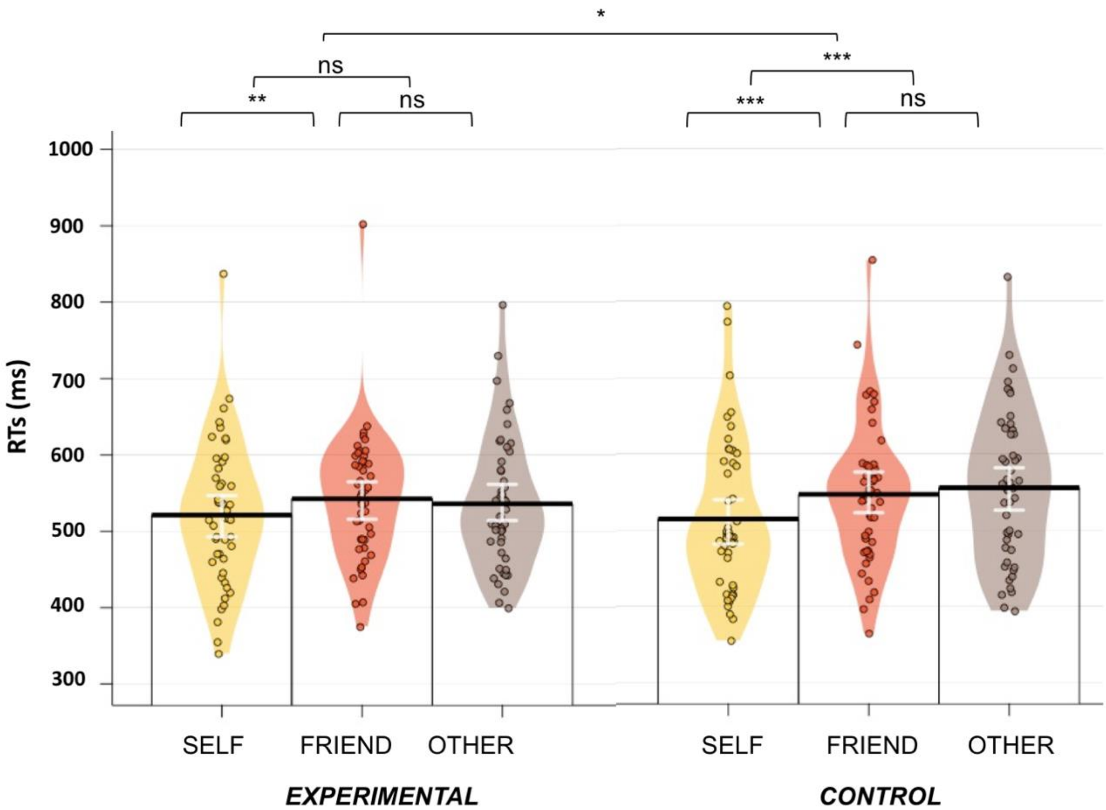

Figure 17. Experiment 7**:** Mean reaction times (RT) as a function of the voice identity (self, friend, or other) and condition (experimental vs. control, in which experimental participants hear their true self-voice presented as 'other'). Graph models match trials only. The error bars indicate the SEs of the means. Coloured segments show smooth density curves for the full data distribution, while individual dots indicate mean RTs per participant. The top bar indicates the interaction, horizontal bars show post-hoc comparisons. Asterisks denote significance within RTs as determined by likelihood ratio tests. *** p<.001

In matched trials, the interaction between *condition* (experimental, control) and *voice identity* (self, friend, other) was significant ($\chi 2(2) = 8.14$, p=.01). Specifically, the new self-voice was prioritised relative to the friend-voice (p<.001) and the other-voice (p<.001) by control participants only. For control participants, all three voices were unfamiliar and arbitrarily assigned to self, friend, and other. In the experimental group by comparison, reaction times to the new self-voice were significantly faster than reaction times to the friend-voice (p=.002) but not to the other-voice (p=.06). Here, the other-voice was actually the participants' true self-voice. In neither group did the reaction times to the friend-voice differ from the other-voice.

These results demonstrate that participants attributed bias to a new self-owned voice relative to a friend-voice even in the presence their true self-voice. However, the bias for a new self-voice was not great enough for it to be significantly prioritised over participants' true self-voice which was presented as belonging to an other. Indeed, the typical bias for the self over an other is here diminished when “other” is actually the true self-voice.

In mismatch trials, the interaction between *condition* and *voice identity* was non-significant ($\chi2(2) = 3.373$, $p=.185$). However, there was a significant main effect of *voice identity* ($\chi2(2) = 18.296$, $p<.001$), in which reaction times to the other-voice were, unusually, significantly faster than reaction times to either the self-voice ($p=.014$) or the friend-voice ($p<.001$) in mismatched trials. This was driven largely by the experimental group, in which there remained a small bias for the true self-voice despite it being temporarily associated with “other” within this task. Finally, the main effect of *condition* was also non-significant ($\chi2(1)=3.21$, $p=.073$).

Sensitivity

The interaction between *condition* and *voice identity* was non-significant ($\chi2(2) = .106$, $p=.948$) as was the main effect of *condition (*$\chi2(1) = 1.229$, $p=.2676$) and the main effect of *voice identity* ($\chi2(2) = 3.303$, $p=.192$). Here then, there is not a self-prioritisation effect in either group and neither group differed significantly in their perceptual sensitivity to each of the voices.

Accuracy

Here the interaction between *condition* and *voice identity* was non-significant in the match trials ($\chi2(2)= 5.47$, $p = .064$), as was the main effect of *condition* ($\chi2(1) = 3.08$, $p=.079$). However, there was a significant main effect of *voice identity* ($\chi2(2)=56.01$, $p<.001$). Accuracy for the self-voice did not differ significantly from accuracy for the friend-voice by either experimental or control participants. Yet accuracy for either voice was greater than accuracy for the other-voice ($p<.001$). This demonstrates that accuracy for the other-voice was considerably reduced in both groups, a result that is typical of the perceptual matching paradigm and here shown to be unaffected by the other-associated stimulus actually being the true self-voice. Given that the main effect of *condition* was non-significant, these data further suggest that experimental participants did not find the task any more cognitively demanding, as their overall accuracy is similar to controls. Full model outputs are reproduced in Supplemental Table 18.

In the mismatch trials, the interaction between *condition* and *voice identity* was non-significant ($\chi2(2)$= .014, p = .992). However, there was a significant main effect of *voice identity* ($\chi2(2)$=8.186, p = .016) and post-hoc comparisons showed that participants' accuracy mirrors the match trials. Again, across groups, accuracy for the self-voice did not differ significantly from accuracy for the friend-voice. However, accuracy was better for both the self-voice relative to the other-voice (p=.036) and for the friend-voice relative to the other-voice (p=.036).

Overall, there was a self-prioritisation effect across measures, with the fastest reaction times and highest accuracy shown in response to the newly associated self-voice in both groups of participants. Importantly, I show this self-prioritisation despite the conflict of the true self-voice being presented as owned by a stranger for the experimental participants. Further, the enhanced bias for the true self-voice evidenced in Experiment 6 was diminished here when the voice was temporarily associated with an other. This suggests that, within this task, participants perceptually prioritised a voice based more on the relevance of the concept it was associated to (i.e. self, friend, or other) rather than on its inherent self-relevance, and indeed salience, outside of this association.

4.3.3.2 Interval Estimation Task

Descriptive statistics for estimates in both active and passive trials, at each interval according to the outcome are given in Table 10. Interval estimates are also plotted in Figures 18a and 18b.

Table 10. *Mean interval estimates (ms) across conditions in Experiment 7.*

	ACTIVE TRIALS				**PASSIVE TRIALS**			
	EXPERIMENTAL		CONTROL		EXPERIMENTAL		CONTROL	
INTERVAL	**Self**	**Other**	**Self**	**Other**	**Self**	**Other**	**Self**	**Other**
100ms	128	169	117	153	169	234	181	237
400ms	281	342	255	323	386	458	358	435
700ms	484	559	476	539	600	669	577	642

To assess sense of agency, I ran LMMs on the interval estimates in which an interaction between *agency, condition*, and *voice identity* was modelled.

E.g. lmer(interval estimation ~ agency * condition * voice identity + 1 | participant + 1 | interval +, REML = FALSE).

The three-way interaction was non-significant, as were all two-way interactions between *agency, voice identity*, and *condition.* However, the main effect of *agency* was significant ($\chi2(1) = 885.99$, $p<.001$) showing that interval estimates were significantly lower in the active trials relative to the passive trials ($p<.001$) for participants across both groups. This shows that participants experienced a greater sense of agency in the condition under which they did, truly, have agency in generating an outcome and that this was the case over both the self-voice outcome and the other-voice outcome. As before, however, this difference in conditions may also be influenced by differences in trial structure.

Further, the main effect of *voice identity* was significant ($\chi2(1) = 400.02$, $p<.001$), with post-hoc comparisons revealing that estimates were significantly reduced for intervals terminating in the self-voice outcome relative to the other-voice outcome ($p<.001$), but this was across both active and passive trials. This suggests that the self-associatedness of a voice does interact with temporal perception but, given judgements were affected in the passive trials, this may be independent of notions of agency. Finally, the main effect of *condition* was non-significant, suggesting that participants who heard their true self-voice presented as belonging to an other experienced equivalent agency and similar temporal compression of intervals before the self-voice outcome as control participants. Full model outputs for this analysis are reproduced in Supplemental Table 19.

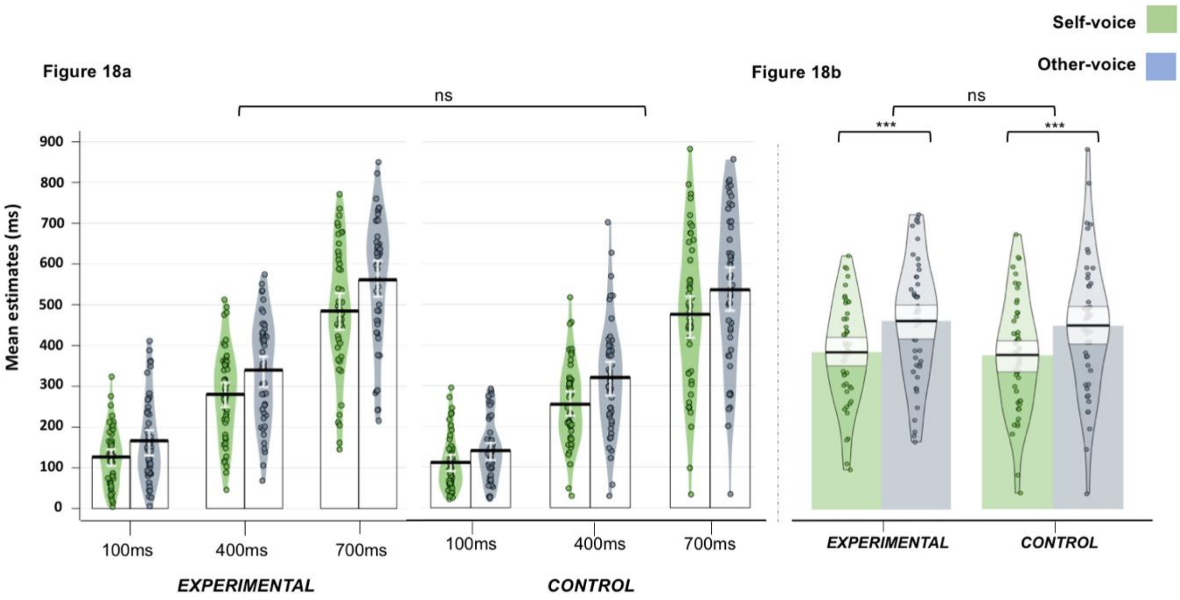

Figure 18a. Experiment 7**:** Mean interval estimates (ms) according to the outcome (self-voice or other-voice), and the condition (experimental vs control). Intervals are plotted according to the true interval duration (100ms, 400ms, and 700ms) and for active trials only. The error bars indicate the SEs of the means. **Figure 18b.** Experiment 7**:** Mean interval estimates (ms) according to the outcome (self-voice or other-voice), and the condition (experimental vs control) Estimates are collapsed across interval duration. Inference bands indicate CIs. Coloured segments show smooth density curves for the full data distribution, while individual dots indicate mean estimates per participant. Top bars show interactions, horizontal bars show post-hoc comparisons. Asterisks denote significance within estimates as determined by likelihood ratio tests. *** p<.001

Overall, these results suggest that participants experienced greater temporal binding over whichever voice has been newly associated with the self, even when the participants' true self-voice was presented within the task but as a vocal outcome associated with an other. These results are discussed further below.

4.4 General Discussion

The true self-voice is a signal that is self-generated, highly, personally familiar and identity-congruent in comparison to a newly self-associated voice which is only self-owned. These factors may make the true self-voice comparatively more self-relevant. In Experiment 6, there was evidence of enhanced bias for the true self-voice; it was prioritised over a friend-voice or an other-voice relatively more than a new self-voice was. Experiment 7 therefore asked whether this enhanced prioritisation was underpinned by the greater relevance of the true self-voice to the self, in comparison to a voice that is only self-owned. Contrary to this hypothesis, Experiment 7 shows for the first time that perceptual prioritisation is influenced by a stimulus' relevance to the self specifically within the constraints of the task. When ownership of the true self-voice was assigned to another, the *enhanced* bias it had previously received in Experiment 6 was diminished. This suggests that by manipulating ownership of the voice and, specifically, by assigning it as other-owned, it made it temporarily less relevant to the self than it would have been if it were assigned as 'self' within the task. Yet, the presence of the true self-voice still impacted participants' performance; participants attributed bias to both their new self-voice and their true self-voice. This disrupted the typical self-prioritisation effect. The results indicated that participants were biased towards their new self-voice and prioritised it over the friend-voice; however, they also prioritised their true-self voice (presented as an other-voice) relatively more than a stimulus associated with an other would typically be. As a result, reaction times did not significantly differ between the voices presented as 'self' and as 'other'.

Experiments 6 and 7 also explored how participants' sense of agency over a self-voice was affected by the degree to which the voice was self-relevant. Importantly, both experiments examined agency over a voice that was self-owned relative to voice that was other-owned. However, in Experiment 6 the self-owned voice was participant's true self-voice while the other-voice was unfamiliar. Conversely, in Experiment 7, it was the self-owned voice that was unfamiliar, while the other-owned voice was actually the true self-voice and, therefore, highly familiar. Strikingly, the pattern of results was equivalent across the two tasks: interval estimates were significantly shorter when they preceded the voice that participants self-owned within the context of the task.

On the one hand then, participants were sensitive to the outcome being self-associated, with significantly lower estimates for intervals terminating in the self-voice outcome relative to the other-voice outcome. On the other hand, they were not sensitive to whether the self-voice outcome was the newly associated voice or the true self-voice. Additionally, their estimates were seemingly unaffected by the true self-voice being presented as belonging to another; estimates were consistently reduced for whichever voice participants had been arbitrarily told belonged to them.

Again, it is important to remember that, in actuality, participants had equal agency in generating all the voice outcomes. For instance, in all active trials, participants made a key-press and heard a voice outcome as a result. The voice outcome was presented without any suggestion of who the voice belonged to, apart from the learnt associations between voices and identities from the previous task. It was only after estimating the interval between their key-press and the vocal outcome that participants judged who the voice had belonged to: self or other (labelled as stranger). Therefore, the fact that interval estimates differed between these voice outcomes – despite equal agency in physically generating them – suggests that it was the experience of the outcome itself that postdictively influenced perception of the interval duration.

The true self-voice should intrinsically signal the self-identity as the agent who caused its occurrence and yet, here, it is consistently the new self-owned voice that garners shorter interval estimates. Moreover, when the true self-voice was presented as temporarily belonging to another, it elicited interval estimates equivalent to that of an other's voice. Indeed, control participants, for whom the voice was genuinely a stranger's voice, gave similar interval estimates as participants for whom it was actually their own voice. Thus, the knowledge that the true self-voice had been temporarily dissociated from self and instead represented an "other" appears to have been more influential on estimates than the cues inherent within the true self-voice. This is supported by previous studies which demonstrate that prior belief can modulate temporal binding (Desantis et al., 2011) and the current results highlight the extent to which prior beliefs about which stimuli are self-owned affect people's judgements on whether they are likely to be self-caused or other-caused.

However, the effect of *voice identity* (i.e., whether the outcome was the self-voice or other-voice) was also evident in passive trials in Experiment 7. Here, again, there were shorter estimates for intervals preceding a self-voice than an other-voice, even when the participants made no action to generate this outcome. Thus, it is unclear what the effect of self-association is modulating here, if not a sense of agency. I had previously hypothesised that the presence of an effect in both active and passive trials could be indicative of self-association influencing estimates *outside* of notions of agency. Specifically, that a self-owned voice may interact with interval estimates because it is perceived and recognised faster. The quicker recognition may affect participants' temporal perception of the preceding interval, without necessarily interacting with agency. Yet, neither Experiment 6 nor 7 support this argument.

In Experiment 6, the effect of *voice identity* was non-significant in the passive trials, suggesting that the effect only influenced agentic trials and thus, that self-association did, truly, modulate sense of agency. In Experiment 7, however, the effect was present in the passive trials but was not of any greater magnitude for the true self-voice relative to the self-owned voice. This is important because if the difference in interval estimates for the self-voice and the other-voice was underpinned, not by agency, but by how quickly they were each perceived and recognised, then it would follow that the quicker a voice is recognised, the shorter the interval should seem. Yet this is not the case. Participants were significantly quicker in recognising the true self-voice relative to a new self-owned voice in the perceptual matching task but this is not reflected in the temporal judgements in the interval estimation task. Notably, the current results are also in keeping with previous temporal binding studies which only use an active condition and do not assess whether the manipulation of interest also affects passive trials. Thus, within the context of this previous literature, I show that participants have a greater sense of agency over whichever voice is perceived as most self-relevant within the task.

Overall, the results from Experiments 6 and 7, taken together, further our understanding of the influence of self-association on bias and agency. First, these results provide new insight on the extent to which the true self-voice is perceptually prioritised relative to others when it is heard aloud. The results demonstrate that prioritisation is sensitive to the factors that may differentiate the true self-voice from a new, self-associated voice; familiarity, identity-congruence, and self-generation. These factors may lead the true self-voice to be more self-relevant and so more prioritised. Importantly though, I also show that this prioritisation is modulated according to what is immediately self-relevant. A highly familiar, self-associated stimulus can still be relatively deprioritised if it is temporarily distanced from the self by becoming owned by another.

These studies have also shown that the temporal binding effect is comparatively unaffected by the factors that distinguish the true self-voice from a new self-voice. Participants experienced more agency over whichever voice outcome they had previously been told they had ownership over. This suggests that temporal binding is modulated according to participants' prior beliefs about whether a stimulus is self-relevant or not. The fact that the interval estimates were not more sensitive to the presence of the highly self-relevant true self-voice is surprising given that a person's voice is unique and, typically, can only signal them as the agent that generated it. However, participants had equivalent agency in generating the true self-voice as the new self-voice and this may be reflected in the results. If so, this suggests that sense of agency is sensitive to whether the outcome is self-relevant or not, rather than *how* self-relevant it is. Moreover, it is possible that when participants usually hear their voice as an outcome to their action it is "in their heads" and it therefore sounds acoustically different to how it was presented in the task via headphones. Thus, it is possible that the self-voice outcome was somewhat distanced from self and was not perceived as any *more* self-associated than any other voice that they only had ownership of. Further studies in which participants actually use their voice within the task may elicit a different pattern of results. Indeed, this is explored in Chapter 6.

Finally, these experiments have provided a benchmark against which we can corroborate previous findings. I have previously concluded that a new self-owned voice had become, at least temporarily, an extension of the self-concept and processed accordingly as a self-relevant stimulus. This is supported by the results here. Specifically, an external voice stimulus can be attributed self-bias enough to compete with a truly self-associated stimulus (the true self-voice) when it is perceived more self-relevant.

Finally, the temporal binding results suggest that participants infer they have had greater agency over sensory outcomes that are perceived relevant to the self within the context. Specifically, a voice stimulus that has become self-relevant through ownership, elicited greater temporal binding than an other-owned voice, even when the other-owned voice was actually their own, highly familiar voice. Thus, prior belief of what outcomes were more self-relevant within the task – and thus, which outcomes they had more control over generating – was perhaps more influential in inferring agency over that outcome than the identity of the outcome itself. To this end then, participants may indeed have a greater sense of agency over a new self-owned voice relative to a voice belonging to an other.

5 Memory advantage for a new self-voice

Abstract

The self is central to our experience; information associated with the self receives greater attention (Shapiro et al., 1997; Dux & Marois, 2009), is prioritised in our perception (Sui et al., 2012; Payne et al., 2020), and a substantial body of literature shows that it is also better remembered (Cunningham et al., 2008; 2011). Chapter 5 extends the examination of how owning a new self-relevant voice biases our perception and experience of that voice. In two further experiments, I ask whether there is evidence of better memory for items that have become relevant to the self by being communicated through the new self-voice. In Experiments 8 and 9, participants are assigned food and household-related items to own by hearing them aloud in either their new self-voice or an other-voice. Items heard in the self-voice become self-owned, while items that are heard in the other-voice become other-owned. In Experiment 8, a surprise test of item memory found no evidence of better memory for self-owned items relative to other-owned items. Experiment 9 comprised a surprise test of source memory (whether an item was self-owned or other-owned) and, similarly, failed to replicate the memory advantage for self-owned items previously reported in the literature. These results are discussed in terms of task design and the conditions required for self-relevance to bias our cognitive processing.

5.1 Introduction

Information that is 'referred' to the self – or is processed in relation to the self – receives a memory advantage in comparison to information related to other people. In memory research, this is known as the self-referential effect. This effect was typically examined by asking participants to evaluate trait words (i.e. trustworthy, creative, courageous) in relation to either themselves or to others (i.e. *Would 'courageous' describe you?, Would 'trustworthy' describe the President?*). In a subsequent test of incidental recall, the trait words that had been encoded self-referentially were better remembered than trait words that had been associated with the other person. The source of this memorial advantage is demonstrated to be the way in which an item is encoded; items encoded to the self receive greater attention and the memory trace is stronger and more robust in comparison to when we encode information in relation to others (Symons and Johnson, 1997; Bentley, Greenaway, Haslam, 2019).

These initial trait-based tasks required participants to explicitly evaluate each trait in reference to the self or the other. However, the memory advantage has since been shown even when stimuli are encoded to the self through more implicit means. For instance, Turk, Cunningham, and Macrae (2008) presented various objects to participants, each represented by pictorial cards. Each object was shown alongside either a picture of the participant's face or a picture of another's face. The authors demonstrated that there was better item recognition in a subsequent old/new judgement task (i.e., have you seen this object before?) if the object had been presented alongside a picture of the self. This may be due to greater attention afforded to the self-face as a stimulus, which then allows for stronger encoding of the associated object. Still, it demonstrates that there is a memory advantage for items that have been incidentally – rather than explicitly – encoded to the self.

A particularly robust means of associating, or referring, a stimulus to the self is by ownership. Specifically, when participants feel a sense of ownership over external objects (Cunningham, Turk, Macdonald, & Macrae, 2008), or even over ideas (Gregg, Mahadevan, and Sedikides, 2017) they are better remembered and viewed more positively than similar objects or ideas belonging to others (Kahneman, Knetsch, &

Thaler, 1991; Huang et al., 2009). This bias is shown even though ownership over these items is imaginary, transient, and induced by the experimental task.

Cunningham et al., (2008) assigned participants ownership over various shopping items by placing items either in their own shopping basket or an adjacent basket belonging to another (see Figure 19). Items were represented by pictorial cards and were colour coded as either blue or red, signalling ownership either by self or by the other. The experimenter held up each card, said the name of the item, and asked participants to move the card into the correct basket according to colour e.g. self-owned blue items were sorted into the self-owned blue basket. Participants were asked to imagine they owned whatever items were in their basket at the end of the experiment.

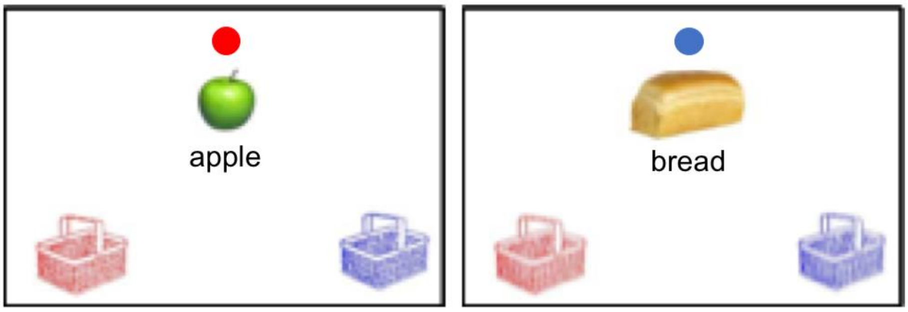

Figure 19. Illustration of the item sorting task (i.e. Cunningham et al., 2008; 2011). Items are sorted into two categories: either self-owned or other-owned, according to the colour coding of the items relative to the baskets which are assigned as belonging to self or other. Items are represented pictorially and also heard auditorily. Figure is reproduced from Cunningham, Brady-Van den Bos and Turk, 2011.

Finally, participants completed a surprise memory test of the shopping items and were shown a digitised pictorial card with the name of the item written on screen. Participants were presented with a mixture of old items and foils and were asked to judge whether they had seen each item before, or not. The results revealed a self-referential effect in which more self-owned items were correctly recognised, relative to other-owned items.

The influence of self-referential encoding has been shown not only in recognition tasks such as these (i.e., testing item memory), but has also been shown to interact with source memory (Lawrence and Chai, 2021; Yin, Ma, Xu, Yang, 2017). That is, memory of the source of the learned information (Pandey, 2011) such as the context in which it was encoded. This is clear from tasks in which participants are not tested on their recognition of whether items have been previously seen or not but, rather, are asked to whom the item was referred. For instance, a study by Cunningham, Brady-Van Den Bos, & Turk (2011) involved self-chosen and self-owned shopping items that could therefore be categorised as either: self-owned and chosen by self, self-owned but chosen by other, other-owned but chosen by self, or other-owned and chosen by other. At test, participants were shown all items again alongside unseen foils. For each item, participants were asked whether the items belonged to the self, the other, or were new. This therefore assessed not only item recognition but, further, source memory; requiring details about the way in which the information was encoded and who it was referred to. Recognition memory was significantly more accurate for the self-owned items that had also been chosen by the self, relative to the other categories. However, there was also a bias towards the self-owned items relative to any other-owned items. This suggests that having ownership of items enables the memory advantage, which may be enhanced if participants have personally chosen to own the items.

The effect of ownership may also underlie the perceptual prioritisation of a new self-voice as presented in Chapter 2; when the voice becomes self-relevant through ownership, processing of that voice is enhanced relative to the voices of others. It is possible, therefore, that the new-voice is not only better recognised but also better remembered, especially after it has been personally chosen. This chapter therefore explores the possibility that through stronger encoding for a new chosen self-voice, items said aloud in that voice are better remembered. The two experiments presented here use adapted versions of the shopping basket task (Cunningham et al. 2008). Experiment 8 tests whether there is enhanced item memory for items that have been assigned to the self via the self-voice. Experiment 9 assesses whether there is enhanced source memory for items assigned via the self-voice.

5.2 Experiment 8

In Experiment 8 I assessed item memory and hypothesised that there would be stronger encoding for items said in the self-voice relative to items said in an other-voice. If so, this would be evidenced by higher accuracy in recognising items presented in the self-voice.

5.2.1 Participants

36 participants (age range = 19-40, mean age = 26.7 years, sd = 6.7 years, 18 female, 18 male) were recruited online via Prolific (www.prolific.ac) as native speakers of English with no visual impairments, hearing difficulties, cognitive impairments or dementia and with an approval rate of over 90%. Participants were then tested online using Gorilla (gorilla.sc, Anwyl-Irvine et al., 2019). Participants were also required to use Google Chrome as their internet browser and to pass a headphone check (Milne et al., 2020) to ensure they were wearing headphones and able to hear the stimuli. Ethical approval was obtained from the Departmental Ethics Committee in Speech, Hearing and Phonetic Sciences at UCL (SHaPS-2019-CM-030). Informed consent was obtained from all participants prior to testing and participants were paid and debriefed upon completion of the study.

5.2.2 Methods

5.2.2.1 Stimuli

The stimuli comprised 216 voice clips saying shopping-related words. This stimulus set size replicates that of the original paradigm (Cunningham et al., 2008). The stimulus set was newly created using synthesised voices that were provided by CereProc Ltd., a company that provides realistic-sounding voices. All 216 words were generated in each of four CereProc voices that were available and then converted to .mp3 files using Audacity (Audacity Team, 2019). Thus, in each voice, there were 216 spoken word clips each naming either a food-related item (i.e., bread, milk, chocolate) or household-related items (i.e. ladle, cushion, clock). These items were then divided into three smaller stimulus sets (Set A, Set B, FOILS; see Appendix for full set) each comprising 72 items such that each set was broadly matched for item type (i.e. food vs household), word length, number of syllables, and word frequency according to the SUBTLEX-UK corpus (van Heuven, Mandera, Keuleers, & Brysbaert, 2014).

5.2.2.2 Procedure

Task 1: Choosing a new voice

Participants were each passively exposed to a voice clip saying '*The house had nine rooms*' from each of the four available synthesised voices. These were presented in randomised order and were included to ensure all available voices were taken into consideration before selection. These voices could then be replayed, and participants were asked to choose one to represent them in the rest of the study *(i.e. 'Which voice do you want as your own?').* Of the four voices available, two were female-sounding voices, two were male-sounding.

After the participants had chosen a voice, they were familiarised with their new self-voice and a voice belonging to an 'other' (labelled as 'stranger'), again via passive exposure. If the participant chose a female-sounding voice for the self, the voice assigned to the stranger was male-sounding. If the participant chose a male-sounding voice, the voice assigned to stranger was female-sounding. This was done to ensure participants could discriminate easily between the two synthesised voices in the later tasks.

To familiarise participants to the voices, each voice was presented and its correct identity label was shown onscreen. Specifically, an identity label of either 'YOU' or 'STRANGER' was displayed in the centre of the screen in black uppercase font on a white background and remained on-screen for 3000ms. 500ms after the label's onset, an auditory exemplar of 'hello' from the correct voice was played. After the auditory stimulus finished, the label remained on the screen until the end of the 3000ms trial. The familiarisation phase consisted of only 4 trials, with each label-voice pairing presented twice in a random order.

Task 2: Shopping Task

In the shopping task, participants were asked to imagine that they were in a supermarket gathering food and household items to buy. Further, that the stranger, whose voice they had just heard, was also there. Participants were then instructed that all shopping items that were available to buy would be said aloud, either in the voice they had chosen for themselves or in the stranger's voice. Importantly, items that were said by their own voice went into their own (imaginary) shopping basket and belonged to them. Likewise, items that were said by the stranger's voice went into the stranger's basket and belonged to the stranger.

Within the task, 144 shopping related words were presented as auditory tokens; 72 were said in the participant's chosen voice – the self-voice – and 72 were said in the stranger's voice. Which set of 72 stimuli was pre-assigned to be self-owned, other-owned (owned by the stranger), or later presented as foils was randomised across participants. Each trial was as follows (see Figure 20); a fixation cross was shown onscreen for 250ms followed by a 100ms blank screen. One of the voices said an item and, at its offset, a screen appeared asking the participants to answer two questions: 1) Did you understand what was said? Yes or No and; 2) Whose item is it? Mine or Theirs.

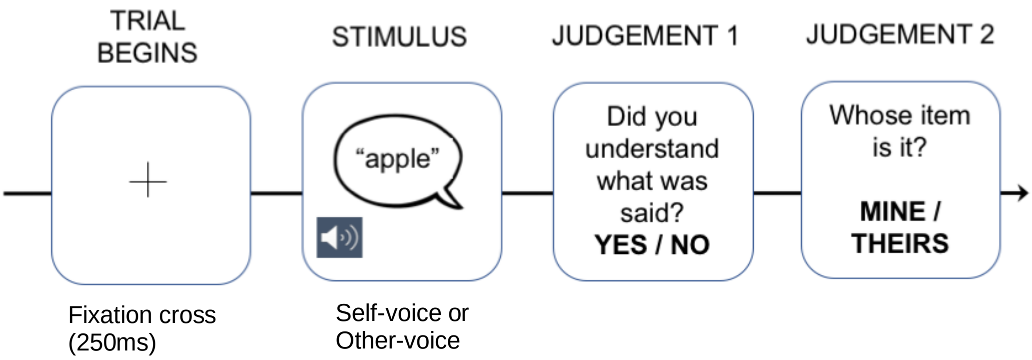

Figure 20. Trial structure for the item sorting task used in Experiment 8. Shopping items are assigned as either self-owned or other-owned according to the voice they are presented in. Voices are associated as belonging to either self or other through familiarisation in Task 1.

The first question was included to ensure that any items that were not recognised in this task were later removed from the analysis of the memory task. Given the use of synthesised voices, some items were more discernible than others according to the chosen voice. Thus, it was important to check comprehension prior to the memory test to ensure that only items that had been understood were tested; there would likely be poor item memory for items that had not been understood. The second question required participants to judge to whom the item now belonged (self- or other-ownership) based on whose voice said the item's name. This measure also allowed us to ensure participants paid attention to distinguishing between the two voices and provided a measure of voice recognition performance accuracy.

Task 3: Memory Task

Participants then completed a surprise memory task. Specifically, I assessed participants' accuracy in recognising whether an item had been presented in the previous task or not, via an old/new judgement task. All 144 of the shopping items were presented again plus an additional 72 foils that participants had not previously heard. Importantly, all 216 words were presented visually onscreen in text, rather than auditorily in one of the synthesised voices. A fixation cross was shown onscreen for 250ms followed by a 100ms blank screen. An item name was then displayed in the centre of the screen in black uppercase font on a white background and remained visible until participants judged it to be either 'old' or 'new' by clicking corresponding buttons on screen.

All 216 items were presented in three blocks of 72 with breaks in between each block. Participants were not provided any feedback in the task.

5.2.2.3 Design

Experiment 8 consisted of one within-subject factor of 'voice identity' (*self,* or *other*) which is here analogous to *ownership*; items were self-owned because they were presented in the self-voice, items were other-owned because they were presented in the other-voice.

5.2.2.4 Pre-processing and exclusion criteria

In the shopping task, participants were asked two questions, which were used to both engage the participant and to inform exclusion criteria. Question 1 asked whether they had understood which item was being referred to in the voice clip. As the study was designed to assess item recognition in the subsequent memory task, any items that each participant reported had not been understood were excluded from the analysis of the memory task. This accounted for less than 0.3% of trials. Question 2 asked participants to identify who the item belonged to, according to which voice said the item. Any items that had been misattributed to either self or other were also excluded as this showed participants failed to distinguish between the two voices and may not have been engaged in the task. This accounted for a further 1.4% of trials.

In line with the previous study (Cunningham et al., 2008), each participant's hit rate for self-owned items and other-owned items was corrected by subtracting the baseline false alarm rate for judging foils as “old” items. Overall, false alarms comprised 17.3% of participants' responses to foils, which is considerably higher than the previous study in which only 7.2% of responses were false alarms (Cunningham et al., 2008). Whole datasets for two participants with false alarm rates over 50% were excluded and not replaced, as this indicated erratic responses.

5.2.2.5 Analysis

To assess whether there was better recognition for the self-owned items (i.e., items assigned to self via the self-voice) or for other-owned items, corrected hit rates for each ownership group were analysed in a generalised linear mixed model (using lme4). The model included *ownership* (self-owned items vs other-owned items) as the only predictor and three random intercepts: participant, stimulus set, and voice (i.e., which of the synthesised voices they heard as self and as other).

E.g. lmer(corrected hit rates ~ ownership +
1 | participant + 1 | stimulus set + 1 voice, REML = FALSE)

Statistical significance of the effects was established via likelihood ratio tests by dropping the fixed effect of ‘ownership’ from a full model that included the random intercepts. For all analyses reported here, post-hoc comparisons were conducted in *emmeans* (Lenth, 2016) and were adjusted for the multiple comparisons via Bonferroni correction.

5.2.3 Results and Discussion

The uncorrected hit rates from Experiment 8 are plotted in Figure 21. The correct hit rates were then analysed. Here, the main effect of *ownership* was non-significant ($\chi2(1) = 0.873$, $p =.351$) This indicated that item memory did not significantly differ for self-owned items or other-owned items.

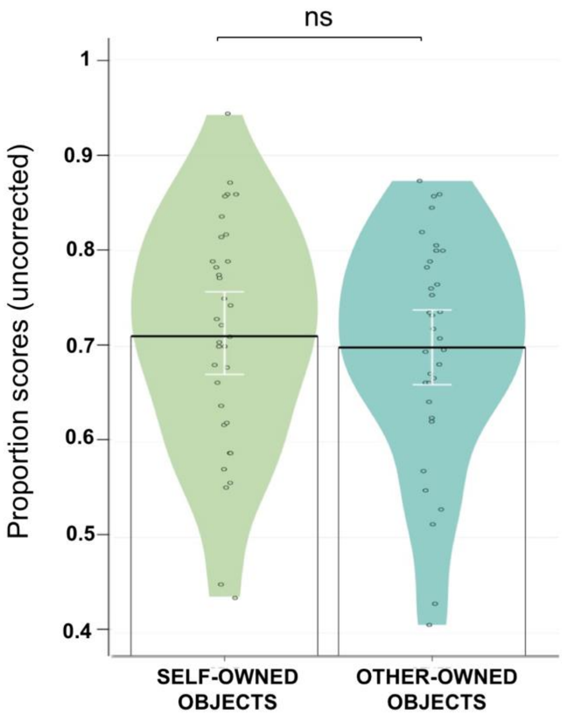

Figure 21. Experiment 8: Proportion scores (uncorrected hit rates) according to ownership (self-owned or other-owned). The error bars indicate the SEs of the means. Coloured segments show smooth density curves for the full data distribution, while individual dots indicate mean scores per participant.

Within these results there was no detectable memory advantage for self-owned items relative to other-owned items. Therefore, the current results suggest that items encoded self-referentially via the self-voice are not better remembered.

It is possible that this null result is because the self-voice was not associated to participants' concept of self strongly enough to trigger self-referential encoding for the objects said in that voice. Previous studies suggest that self-referential encoding can occur both when the objects are explicitly referred to self (i.e., traits evaluated in reference to self) and, as here, implicitly referred by association to self (i.e. objects better remembered when presented alongside the self-face). Here then, I hypothesised that hearing items said aloud in the self-voice would implicitly enable self-referential encoding such that those items were processed in relation to the self.

This was not the case. It is possible then that the relation was too implicit; that the item being presented in a new self-voice was not equivalent to items being presented alongside the true self-face. The additional distance, from self, to self-associated voice, to self-owned object was too weak a link for participants to implicitly assume self-ownership over items said in that voice and, thus, items were not encoded self-referentially. Indeed, the memory advantage for ‘self’ tends to be larger when the stimulus is encoded explicitly (Turk, Cunningham, and Macrae, 2008).

Given that the effect of self-referential encoding for self-owned items is, typically, robust, the lack of memory advantage here may also be due to the task design. It is possible that the task was less engaging when run online, remotely, compared to previous iterations of this task. In a similar online study by Sparks, Cunningham and Kritikos (2016), the memory advantage was influenced by the degree to which participants actively engaged with objects in the encoding phase. When participants were able to sort items by a quick keypress as being either self- or other- owned according to the colour coding cue, it did not necessitate maximum attention on the object itself. Indeed, a memory advantage for self-owned items was only detected when participants were encouraged to focus more on the objects at the encoding phase. In the original study, ownership was also assigned by colour-coded stickers placed on each pictorial representation of an image. However, the experimenters placed these coloured stickers in different locations on each image to ensure participants focused on every object in turn.

It was not possible to replicate this equivalently in the current study using only auditory stimuli; thus, I instead included a judgement of whether the item was discernible and, second, in which voice it was said. It is also possible that this task design actually split participants’ attention by asking whether they were able to discriminate the item in each auditory clip. That is, visually recognising an object and an associated colour-cue as in the original task may engage very similar underlying processes. In contrast, requiring participants to complete an auditory speaker recognition task simultaneously to object (word) recognition may have engaged different mechanisms. By splitting participants’ attention, the memory advantage may have been reduced. Turk, Brady-van den Bos, Collard, Gillespie-Smith, Conway et al. (2013) demonstrate that the typical effect of self-ownership on memory only arises when sufficient attentional resources are available. If this task was too demanding on attentional resources at the time of encoding, any memory enhancement through self-referential encoding may have been diminished.

Relatedly, the fact that participants were presented the items only auditorily rather than visually may have increased the difficulty of the task. It is widely established that items presented as pictures are better remembered - with higher fidelity – than the same item presented auditorily (Ally & Budson, 2007; Bigelow & Poremba, 2014; Cohen, Horowitz, Wolfe, 2009; Grady, McIntosh, Rajah, Craik, 1998). The auditory stimuli were only presented for the duration it took for them to be heard. In contrast, the original task required participants to physically sort pictorial cards into baskets which may have maximised engagement with each object. Indeed, performance in the current experiment was ~10-15% lower than in previous studies when comparing corrected hit rates (i.e., Cunningham et al., 2008) and this was due to a higher false alarm rate. This may be underpinned by data quality; poor performance accuracy may also reflect low engagement with the task. Future iterations of this task should work to better assess attention and direct focus to each item for longer.

Overall, in contrast to previous literature, the current study shows that item memory is not significantly better for self-owned items relative to other-owned items. Here, this is analogous to saying that there is no better memory for items because they have been encoded to the self via a self-voice. Experiment 9 expanded this work and aimed to address a possible advantage for source memory, if not item memory.

5.3 Experiment 9

In Experiment 9, I assessed source memory and hypothesised that there would be enhanced source memory for items encoded as self-owned than other-owned. If so, this would be evidenced by higher accuracy in identifying the ownership of self-owned items than other owned items. Here, again, items became self-owned implicitly, signalled by the self-voice.

5.3.1 Participants

36 participants (age range = 18-40, mean age = 28.1 years, sd = 6.4 years, 18 female, 18 male) were recruited online via Prolific (www.prolific.ac) as native speakers of English with no visual impairments, hearing difficulties, cognitive impairments or dementia and with an approval rate of over 90%. Eligible participants were then tested online using Gorilla (gorilla.sc, Anwyl- Irvine et al., 2019). As before, participants were also required to use Google Chrome as their internet browser and to pass a headphone check (Milne et al., 2020) to ensure they were wearing headphones and able to hear the stimuli. Ethical approval was obtained from the Departmental Ethics Committee in Speech, Hearing and Phonetic Sciences at UCL (SHaPS-2019-CM-030). Informed consent was obtained from all participants prior to testing and participants were paid and debriefed upon completion of the study.

5.3.2 Methods

5.3.2.1 Procedure

Participants completed the same three tasks as in Experiment 8, including choosing a new voice, the shopping task and finally a memory task. Experiment 9 used the same stimuli as in Experiment 8. However, the tasks were adjusted as follows. First, in the shopping task, the question asking whether or not participants understood the auditory presentation of the item was removed. In Experiment 8, negative responses counted for <0.3% of trials and the task demands may have split participants' attention. Thus, in Experiment 9, the only question participants answered in the shopping task was: Who does the item belong to? As before, successfully answering this question required identifying which of the two voices – self, or other – the item was said in.

Second, in the memory task, I wanted to test source memory rather than item recognition. As such, the old/new judgement task included in Experiment 8 to test item memory was removed. Participants were instead asked to report whether each shopping item was owned by Self, Other, or Neither (in the case of foils). Every other part of the procedure remained the same.

5.3.2.2 Design

Experiment 9 consisted of one within-subjects factor of 'voice identity' (*self, other*) which, again, is analogous to 'ownership: it was self-owned if it was said in the self-voice, other-owned if it was said in the other-voice.

5.3.2.3 Pre-processing

As before, participants' responses in the shopping task informed the exclusion criteria. Any shopping items for which participants misidentified the correct owner (according to which voice said the item) were excluded. Erroneous responses indicated that they failed to discriminate between the two voices and/or did not correctly encode the item either as self-owned or other-owned. This counted for <1% of trials.

For each participant, unbiased hit rates (H_u scores; Wagner, 1993) were calculated to more precisely model participant's accuracy in judging items to be either self-owned, other-owned or foils (i.e., previously unheard, unowned items). H_u scores provide a measure of accuracy according to both stimulus frequency and response frequency. That is, H_u scores combine the probability that a stimulus is a) correctly identified according to how many times it has been presented and b) that a particular response is correctly used according to how many times that response is appropriate.

For example, for all the self-owned items, instances in which the item was correctly identified as being self-owned were coded as *hits.* Instances in which the item was judged to be self-owned when it was actually other-owned or a foil were coded as *false alarms.* Instances in which the item was incorrectly judged to be other-owned or a foil when it was actually self-owned were coded as *misses.* Instances in which the item was correctly judged to be other-owned or a foil were coded as *correct rejections* (i.e., participants correctly rejected the item as not self-owned). A confusion matrix was thereby created per participant (see Figure 22a) and, from this matrix, the calculation for unbiased hit rates (i.e., H_u scores) could be applied (see Figure 22b). All H_u scores were arcsine transformed (Wagner, 1993) and the same coding process was then repeated for the two other response categories: other-owned and foils. To compare this performance with chance, an individual chance level was also calculated per participant, per condition (Wagner, 1993). Chance was calculated as the joint probability that a stimulus would co-occur with a response of a corresponding category (i.e., a hit) by multiplying together the independent probabilities of each of these occurring alone (see Figure 22c).

22a

RESPONSE	SELF-OWNED	OTHER-OWNED	FOIL	Total
Self-owned	(a) 43	(b) 8	(c) 21	**72**
Other-owned	(d) 4	(e) 41	(f) 27	**72**
Foils	(g) 2	(h) 3	(i) 67	**72**
Total	**49**	**52**	**115**	

22b

$$H_u = \frac{a}{(a+b+c)} \times \frac{a}{(a+d+g)}$$

22c

$$\text{Chance} = \frac{(a+b+c)}{N} \times \frac{(a+d+g)}{N}$$

Figures 22a-c. 22a: An example confusion matrix for one participant in Experiment 9. **22b:** The calculation applied to the confusion matrix to calculate Hu scores (unbiased hit rates) per participant, per condition. **22c:** The calculation applied to the confusion matrix to calculate chance level performance per participant, per condition (see Wagner 1993).

5.3.2.4 Analysis

Participants' arcsine-transformed, unbiased hit rates (H_u scores) for self-owned items and other-owned times were assessed in a linear mixed model (using lme4). The model included *ownership* (self-owned items vs other-owned items) as the only predictor and three random intercepts: *participant, stimulus set*, and *voices* (i.e. which of the synthesised voices they heard as self and as other).

E.g. lmer(unbiased hit rates ~ ownership + 1 | participant + 1 | stimulus set + 1 | voice, REML = FALSE)

Statistical significance of the effects was established via likelihood ratio tests by dropping the fixed effect of 'ownership' from a full model that also included the random intercepts.

5.3.3 Results and Discussion

The H_u scores from Experiment 9 are plotted in Figure 23. Here, the main effect of *ownership* was non-significant ($\chi^2(1) = 1.477$, $p = .22$) and this indicated that source memory did not significantly differ for self-owned items or other-owned items.

Thus, there was no evidence of enhanced source memory for self-owned items relative to other-owned items. This suggests that source memory was not enhanced for items because they were referred to self via the self-voice. This result tallies with Experiment 8 such that I, overall, fail to replicate the typically advantageous effect of self-referential encoding on either item memory or source memory.

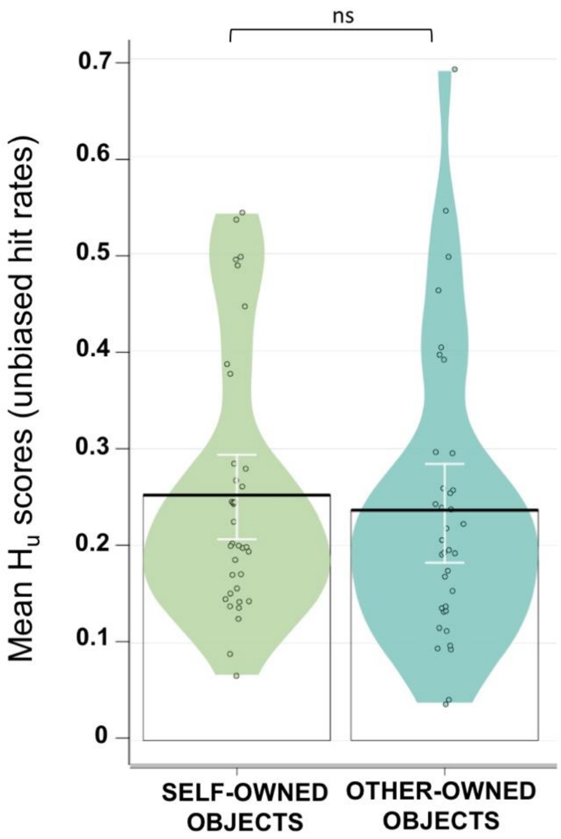

Figure 23. Experiment 9: Mean H_u scores (unbiased hit rates) according to ownership (self-owned or other-owned). The error bars indicate the SEs of the means. Coloured segments show smooth density curves for the full data distribution, while individual dots indicate mean scores per participant.

As in Experiment 8, it was necessary to consider how memory may have been affected by the level of engagement with the items at the encoding phase (i.e., in the shopping task). In this task, participants were only required to judge who the item belonged to, according to the voice in which it was said. It is possible that this could be correctly determined only by a process of discriminating between the two voices, without giving due attention to the content of the voices (i.e., the items themselves). Thus, the lack of memory enhancement for the self-owned items here may again be because the items did not receive sufficient attention for self-referential encoding. Performance accuracy here was slightly lower relative to previous studies assessing source memory; raw accuracy was at 71% for identifying objects that were self-owned compared to ~82% in Allan et al., 2017, for instance). However, it is clear from the current results that participants were performing above chance (Figure 24).

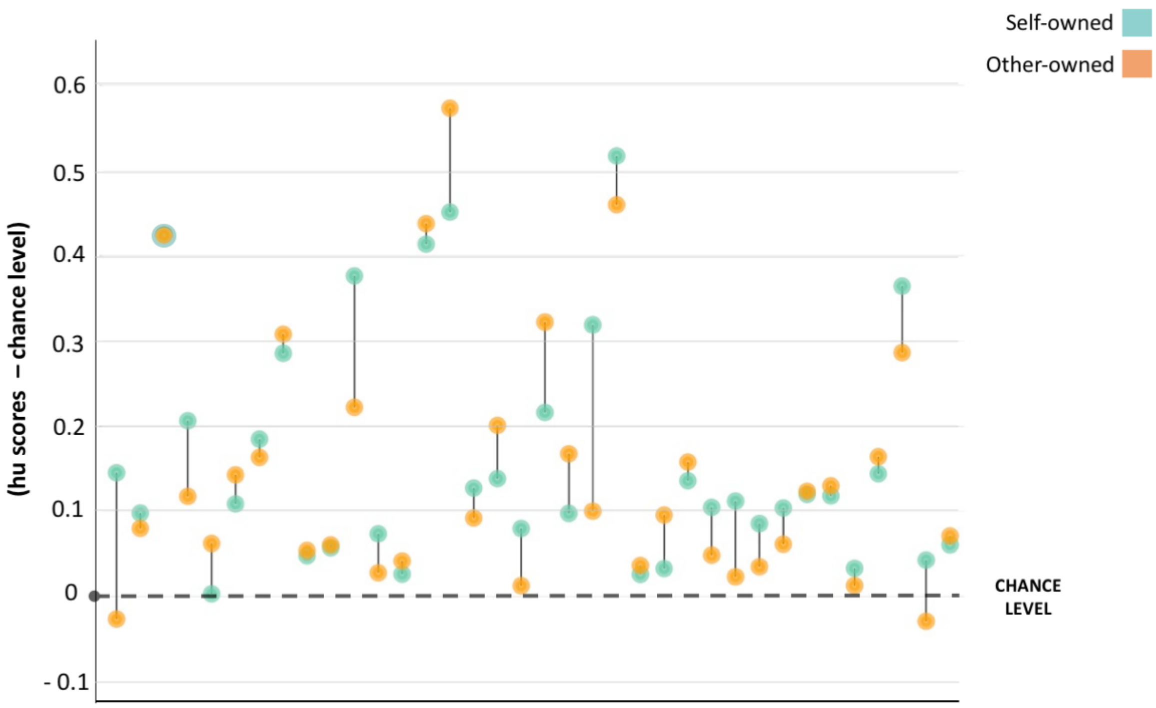

Figure 24. Experiment 9: Participants' individual performance accuracy (Hu scores i.e., unbiased hit rates, arcsine transformed) in each condition (self-owned objects, other-owned objects) relative to chance level performance for that participant, in that condition. Each vertical line represents an individual participant, with green circular dots representing Hu scores for self-owned objects, orange circular dots for other-owned objects. Positive integers show performance above chance level, negative integers show performance below chance level.

Paired samples t-tests were run to statistically test whether participants performed above chance in this task. Specifically, I compared H_u scores to chance level for both self-owned items and then other-owned items. The t-test comparing participants' performance for self-owned items relative to chance performance for that condition was significant ($t(35) = -4.80$, $p < .001$). This shows that participants' memory performance was significantly better than chance for self-owned items. Similarly, the t-test assessing the difference in performance for other-owned items relative to chance was also significant ($t(35) = -6.07$, $p<.001$). Thus, all participants were performing above chance, suggesting reasonable engagement with the task. Despite this, there was no evidence of a memory advantage for self-owned items or, therefore, for items referred to the self via the self-voice.

5.4 General Discussion

Across Experiments 8 & 9, I assessed memory for self-owned items that had been assigned to the self implicitly via their association with the self-voice. Within the current task design, there was no evidence for a memory bias for either the self-voice or the self-owned items.

It is clear from the literature that there is typically better memory for self-owned items because they are associated with the self. Relatedly, from the studies presented within this thesis, it is clear that there is better recognition of a new self-voice - again, because it is associated with, and owned by, the self. These associations are therefore direct; the saliency of a new stimulus can be altered because of its direct relevance to the self-concept. Within the current task design then, it is possible that there was no bias for the self-owned items because they did not become associated with the self but, rather, with the self-voice. That is, the task essentially required a two-step process of referring stimuli to the self; first between the self and a new voice and then between the new voice and the self-owned items. The self-owned items could only become self-relevant by the fact that they were associated with the self-voice. Further studies may serve to test whether this additional step affects the strength of self-association and therefore the attribution of self-bias.

Given the lack of effect here, it is interesting to consider the robust effect of self-bias for a self-voice in speeded judgement tasks, such as in the perceptual matching paradigm. Importantly, this task also relies – implicitly – on the memory of what has previously been encoded self-referentially (i.e., has been associated with the self) from what has been encoded other-referentially (i.e., has been associated with other). Yet, in Experiments 8 and 9, the typical bias for the self-voice does not extend to more accurate memory for the *content* of what was said in that voice (i.e., the items themselves). This suggests that the processes by which self-referential encoding enhances our processing may differ according to the task demands, be it faster perceptual processing or stronger memory traces.

However, it is also possible that the lack of effect was because of the lack of engagement in the task; self-referential encoding relies on sufficient attention being given to the stimulus at the time of encoding. Further, the typical memory advantage relies on conscious awareness of the information being self-relevant (Kim, Jeon, Banquer, & Rothschild, 2018) which may not have been stressed within the task. Although the voices were understood as self-relevant, it is possible that the relevance of the items themselves was not emphasised explicitly enough. Future iterations will need to devise ways of maximising attention, emphasising what is self-relevant, and assessing engagement with the items. This could involve requiring participants to transcribe the heard word, and/or presenting the item both pictorially and auditorily. Further, it could be worthwhile to include the self-voice and other-voice only as a means of assigning ownership over items (e.g., saying 'mine' where relevant, akin to colour coding) rather than as a means of presenting the item as well. Additionally, when this task is conducted in person, the participant has a physical representation of an 'other' and the objects that the other owns (which, are therefore not self-associated). It may be that participants did not distinguish between self and other within this task because self-ownership over the items was so arbitrary, and the other identity was essentially absent. This could have reduced the extent to which self-owned objects were encoded differently from other-owned objects.

Lastly, this implementation of the shopping basket task was limited in its ability to elucidate a memory bias for self-owned items from bias for the self-voice. That is, if there had been enhanced memory for the self-owned items, it would not have been clear whether this was because the items were self-owned or because they had been said in the self-voice. A further study could disentangle these by presenting each item in either the self-voice or the other-voice but embedded in a sentence such as '*The apple is mine*', or '*The apple is not mine*'. This would allow for participants to be assigned ownership over items by both the self-voice and other-voice and would allow a contrast between memory bias for items that are a) said in the self-voice and are self-owned; b) said in the self-voice but are other-owned; c) said in the other-voice but are self-owned; d) said in the other-voice and are other-owned. This would provide further insights into how self-bias is influenced by the effect of ownership but also by the way in which it is referred to the self.

Overall, the experiments presented in this chapter suggest that bias for the self-voice does not extend to better memory for items said in that voice, nor for the ownership of those items. While it is possible that this result is underpinned by the task design, it does highlight the sensitivity of self-referential encoding in terms of attentional resources and the salience of items as being self-relevant within the task. Biasing of perceptual processing, by contrast, is considerably more robust and sensitive to the status of sensory stimuli as self-relevant or not. These results raise further questions and provide a starting point for further examination of the boundaries of self-bias and the extent to which a new self-voice can bias cognitive processing purely by it becoming self-relevant.

6 Using a new self-voice in a social context

Abstract

This thesis has explored the roles of ownership and choice in influencing whether – and the extent to which – a new voice is perceived as part of the self. Yet voices are fundamentally social stimuli, and their relevance to the self may be underpinned by the extent to which they can be used to communicate with others. A voice may become better incorporated into the self-concept if it can physically represent the self through its use. In Experiment 10, I assessed the importance of having agency and flexible control over a new voice on how far it was deemed self-relevant. To enable participants to use a new self-voice, a novel two-player online game was created, in which participants were able to use a new synthesised voice to represent themselves and interact with another participant in a 30-minute game. At test, I compared participants who had played the online game to a control group (n=44) who had only brief exposure to the voices. The potential effect of having agency in using the voice was again assessed according to whether the self-voice was perceptually prioritised to a greater extent after it had been self-produced, relative to a voice that was only owned. Relatedly, I assessed sense of agency over a voice that participants had been able to exert control over, relative to a voice that was only self-owned. The results indicated that both experimental participants and control participants perceptually prioritised their self-voice and experienced a sense of agency over it. Critically, having agency over the voice in a social interaction did not modulate bias towards it nor sense of agency over it in the subsequent interval estimation task. Together, the results suggest that the fundamental knowledge of what is “mine” may be sufficient to generate a perceptual bias and a sense of agency, which speaks to the automaticity of this cognitive self-bias.

6.1 Introduction

A voice is 'revelatory of self, mental states, and consciousness' (Sidtis and Kreiman, 2012, p.4). However, a voice cannot be revelatory in this way without vocal action from the speaker. Rather, a voice can only exist if someone has actively produced a sound (Scott and McGettigan, 2015). Unlike faces, which transmit identity cues continually and perhaps even if the person does not want to be identified, it is only through the speaker's *use* of their voice that the voice can represent the self. Thus, it is not through simply having a voice that we achieve our social and communicative goals, but through how we use it.

This is particularly evident for people who are not always able to use their own self-generated voice but rely on an Augmentative and Alternative Communication (AAC) device to communicate. The relative difficulty of using technology to communicate quickly - and with as much nuance as natural speech - can sometimes result in the misperception that 'less talk means fewer ideas' (Wickenden, 2011, p.3). Thus, the ability to use a voice to meet both communicative and social goals is critical to the accurate construction and negotiation of the self-concept in relation to others.

It is important, therefore, to assess the experience of the self-voice, not just as something that is owned but as something that is used. Thus far, this thesis has investigated the self-bias that is attributed to a new self-voice and the agency people experience over it after simply being given, or choosing, the voice as a social stimulus to own. In all cases, participants have only heard their new self-voice and have not actively used it in a social context, either as a means of self-representation or as a means of achieving a communicative goal. This chapter therefore extends these studies to evaluate whether the self-bias attributed to a new voice, and/or the sense of agency over it, is modulated by self-producing the voice in a social context. A novel online environment was created to allow pairs of participants to interact together using new self-voices via text-to-speech technology. In line with the previous chapters, self-bias and agency were subsequently assessed using a perceptual matching paradigm (Sui et al., 2012) and an interval estimation task (Engbert, et al., 2007), respectively.

6.1.1 Bias for self-produced outcomes

Previous studies have shown that the outcomes of self-produced actions are better recognised than those of others' actions. For instance, people can recognise their own biological motion better than the motion of other people (Wolff, 1931; Beardsworth and Buckner, 1981). That is, when asked to make judgements of whether a movement has been self-produced or other-produced, self-produced motion is more accurately recognised. This extends to better recognition of self-produced drawings (Knoblich and Prinz, 2001); handwritten strokes (Knoblich, Seigerschmidt, Flach, & Prinz, 2002); and even self-produced sounds such as clapping (Repp,1987). Repp and Knoblich (2004) have also demonstrated that pianists can better recognise their own performances from those of others who perform the same score. Interestingly, the pianists also rated their own performances more highly on average than those of other pianists and, further, higher than anyone else rated theirs. This is analogous to people rating their own voice both as more attractive than other peoples' and also as more attractive than other people rated theirs (Hughes and Harrison, 2013). The higher ratings for the self-voice in the latter study are assumed to arise because of an implicit egotism - a self-serving bias - which also seems applicable to the pianists' higher than average self-ratings.

The fact that people judge the outcomes they have self-produced more positively is evident in several other studies. For instance, when people are actively involved in self-preparing and producing their own food, it increases their liking of it (Radtke et al. 2019; Dohle, Rall, & Siegriest, 2016). This finding is also evident in the classic IKEA effect (Norton, Mochon, & Ariely, 2012), in which objects that have been self-produced are more highly valued. For instance, when participants are asked to make origami figures (Mochon, Norton, and Ariely, 2012; Norton et al. 2012); furniture items (Bendapudi and Leone, 2003); and customised clothing (Franke et al., 2010), the end product is consistently rated as more valuable than an identical item made by someone else.

Within these self-production studies, the increased value of self-produced items is most prominently seen to be because of the creation of an end product which can signal a competent self-identity to both the self and to others (Mochon, Norton, Ariely, 2012; Norton et al., 2012). Further, because self-produced items can allow individualisation, they also afford greater self-expression (Kaiser et al., 2017; Yoo and Park, 2016). Therefore, the self that is reflected by a self-produced item is more identity-congruent than the self that may be reflected by a generic item. It is possible, however, that such items are also viewed more positively because they have become self-associated through the process of making.

The advantageous effect of self-produced outcomes is also evident in memory. A wealth of research has shown that words are better remembered when they are said aloud, as opposed to silently read (MacLeod et al., 2010) or heard in another's voice (MacLeod 2011; Mama and Icht, 2016). This effect – known as the production effect (MacLeod et al., 2010) – is further increased when words are either read loudly or sung (Quinlan and Taylor, 2013). This finding is presumed to arise because of the increased distinctiveness of the production, which, in turn, imbues a greater memory advantage. Furthermore, this memory advantage remains when participants hear the words they self-produced in an earlier study played back on a recording later (Maslowski, Meyer & Bosker, 2018). This suggests that the act of self-producing outcomes imbues a lasting advantage in memory for those outcomes.

6.1.2 Effects of self-production on perceptual prioritisation

Given the findings reviewed above, there are several factors that suggest that the *use* of a new self-voice may affect the way people perceive and experience that voice. The current study therefore investigates whether producing a new self-owned voice influences perceptual prioritisation of that voice, or sense of agency over it.

First, bias may be increased as the voice shifts from being a self-owned stimulus to one that allows self-expression. Given that the voice, when self-produced, signals the self-identity, it may be deemed more self-relevant than when that signal is only self-owned. This greater self-relevance may enhance bias towards it. Indeed, in Chapter 4, I demonstrated that the true self-voice was prioritised more than a voice that was only self-owned. This greater prioritisation may have been underpinned by factors that make it more self-relevant, such as its being self-produced, and/or more familiar, and/or more identity-relevant.

If people are biased towards self-produced outcomes because they reflect a competent self-identity (Mochon, Norton, Ariely, 2012; Norton et al., 2012), the use of a new self-voice may similarly signal competence to another participant. Although the voice is not a tangible object created manually as in the studies reviewed above, it is through the self-production of vocal outcomes that a communicative message is "made" – the voice allows a talker's thoughts, ideas, preferences and, indeed, social traits such as competence, to be conveyed to another speaker. The ability to use the voice therefore – relative to simply owning a voice - may increase bias towards it if it can be used to signal a competent identity to another.

Relatedly, in Chapter 2, I examined whether a voice that was more identity-relevant with the participant, specifically in terms of gender, could influence bias towards the voice. Identity-relevance here may be analogous to greater self-expression; by gender-matching the voice to the participant, it allowed a more accurate means of self-representation. However, the results suggested that a new voice's similarity to the true self did not affect how much it was prioritised. Thus, I concluded that a stimulus is prioritised according to whether it is self-relevant or not, rather than by the degree to which it is self-relevant. In this previous study though, the voice was not used as a means to represent the self to others and thus, the degree to which that voice was identity-relevant and able to accurately represent the self may have been irrelevant. Here then, once the new voice has become not only self-owned but a tool that is used to interact with another participants, its potential to more accurately represent the self may enhance the bias attributed to it.

Finally, through the use of the self-voice, people should become more familiar with that voice. Holmes, To, & Johnsrude (2021) have recently shown that only 10 minutes of familiarisation to a new voice is sufficient for people to recognise that voice and for its intelligibility to be increased relative to a new unfamiliar voice. Additionally, earlier studies presented within this thesis (Chapter 4) have shown that the true self-voice receives greater perceptual prioritisation relative to a newly self-associated voice. This result may be because the true self-voice is more familiar relative to a new self-owned voice. Taken together, it is reasonable to expect that as familiarity with a new self-voice increases and as the voice is experienced as a self-produced outcome, the degree of prioritisation for it may also increase.

6.1.3 Effects of self-production on sense of agency

The current study also aimed to investigate whether people experience a greater sense of agency over the new self-voice after social use. Typically, temporal binding studies investigate agency over self-generated actions relative to either involuntary actions or no action. Thus, it may be considered that all such binding studies examine agency over self-produced outcomes. In the temporal binding studies presented earlier within this thesis, I assessed the sense of agency over a self-voice outcome relative to an other-voice outcome that was generated within the task. In these experiments the self-voice was differentiated from the other-voice by the participants' ownership of it, but both voices were self-generated in the paradigm. In the current study, however, participants had the opportunity to produce the self-voice in a social setting and gain prior experience of hearing it as an outcome of their actions. Moreover, their agency over the new self-voice was manifested as keypresses to produce typed messages as speech and, similarly, it is also through keypresses that the voice outcomes are generated in the interval estimation task. Thus, the relevance of the new self-voice voice to the self may be further increased subsequent to prior social use, particularly given the common method of production. Here then, the self-voice and the other-voice are further differentiated: the self-voice is not only a stimulus that is owned by participants, but it is one they have self-produced. Intuitively, giving people control, flexibility and, indeed, agency, over using a new voice to represent themselves should imbue a greater sense of agency over that voice relative to people who simply own the voice but cannot use it. The current study tested this hypothesis.

Here then, I assess participants' perceptual prioritisation of a self-voice and their sense of agency over it after it has been self-produced in a socially interactive context in comparison to participants who have not self-produced the voice.

6.2 Experiment 10

6.2.1 Hypotheses

I hypothesised that participants would show increased bias towards the self-voice after self-producing it in a social context (experimental group), relative to participants who had not experienced social use of the voice (control group). I predicted that this result would be indicated by quicker reaction times and higher accuracy in response to the self-voice by the experimental group relative to a control group.

Second, I hypothesised that participants would have a greater sense of agency over the self-voice after producing it in a social context relative to participants who had not experienced social use of the voice. I predicted that this result would be indicated by significantly shorter interval estimates for action-outcomes involving the self-voice relative to the other-voice by the experimental group relative to a control group and in active trials only (i.e., under conditions of agency).

6.2.2 Participants

88 participants completed the study. 44 of these participants were randomly assigned to an experimental group (mean age = 27.1 years, SD = 6.54 years, age range = 18–40, 22 female, 22 male) while the other 44 constituted a control group (mean age = 25.4 years, SD = 6.27 years, age range = 18–39, 22 female, 22 male). This data set arose following the exclusion and replacement of 9 participants based on exclusion criteria outlined below.

All participants were recruited online via Prolific (www.prolific.ac) as native speakers of English with no hearing difficulties or visual impairment and were tested online using Gorilla (gorilla.sc, Anwyl-Irvine et al., 2019). Participants were also required to use Google Chrome as their internet browser, to pass a headphone check (Milne et al. 2019) and have an approval rate of over 90% on Prolific.

None of the participants had taken part in any of the pilot studies or previous studies within this thesis and, upon completion of the study, were paid for their participation. Ethical approval was obtained from the Departmental Ethics Committee in Speech, Hearing and Phonetic Sciences at UCL (SHaPS-2019-CM-030), and informed consent was obtained from all participants prior to testing.

6.2.3 Methods

6.2.3.1 Procedure

Task 1: Choosing a voice

All participants were invited to choose 1 of 11 synthesised voices provided by CereProc (www.cereproc.com) to be their new voice and to represent them in the rest of the study. Synthesised voices were used because it was essential that participants were given extensive control and flexibility in what they chose to say with their new self-voice. This could be achieved most readily with state of the art, text-to-speech synthesis.

First participants were passively exposed to each of the 11 voices in a random order (each saying: "*The house had nine rooms"*). This was to ensure participants gave due consideration to each available voice. Second, participants were invited to replay the voices at their leisure (each voice here said: "*They're buying some bread"*). Available voices included a range of accents (broadly: English, Scottish, American (US)) and both female-sounding voices and male-sounding voices. After selection, all of the remaining voices that had been rejected as a self-voice were discarded and did not feature in the later tasks for that participant.

All participants were then asked briefly about their voice selection. Specifically, they were asked whether they had chosen a new self-voice that aligned with their gender-identity and/or their accent.

Task 2: "Drawing Conclusions" Game

Of the 88 participants that were recruited to the study, 50% of them were randomly assigned to the experimental condition. Experimental participants only (n=44) were given the opportunity to use their newly chosen self-voice in a socially interactive game.

The game took the form of a drawing game, introduced to the participants as being called "Drawing Conclusions" (see https://github.com/wallscope-research/drawing-woz for more information). The game was specifically created so that pairs of participants could interact with one another on a custom designed platform. The game was created in collaboration with researchers in Heriot-Watt University's Interaction Lab and was written by Angus Addlesee in JavaScript, making use of node.js (see https://github.com/wallscope-research/drawing-woz for more information). This enabled a web application to be built which could allow real-time interaction between participants and, particularly, interaction using new self-owned synthesised voices. Players were able to choose and use synthesised voices from CereProc and share a 'drawing canvas' (see Figure 26, though see also 27-29 below) through Google Chrome.

Within the game, participants were assigned roles, one starting as a *narrator*, the other as an *artist.* Both participants were supplied with a separate, and different, picture deck containing 20 images (Figure 25) based on a pre-existing game "Up a Bit" (Cheatwell Games, 2004)

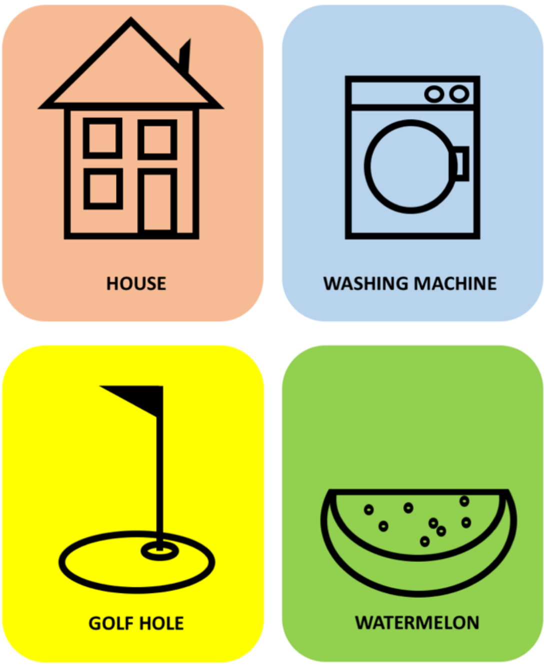

Figure 25. Example images from the picture deck supplied to each participant in the game "Drawing Conclusions".

The narrator chose an object to draw from their picture deck and had to instruct the other participant to draw this object without telling them the name of the object. These instructions were communicated via text-to-speech technology. Specifically, the narrator typed in an instruction, pressed submit, and then the instruction was said aloud in the participant's chosen voice (e.g., *Please start by drawing a square in the middle of the screen*; Figure 26).

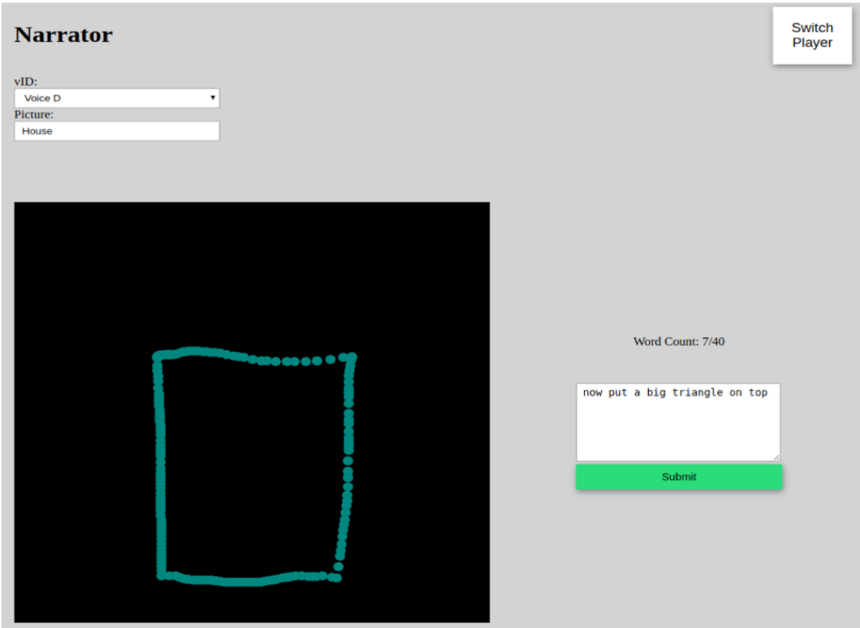

Figure 26. The participants' display when acting as the 'Narrator' within the game, "Drawing conclusions".

Importantly, each participant heard the playback of their self-voice as it was simultaneously transmitted to the other player – the artist – to hear. The artist could then draw the object incrementally as per the narrator's instructions on the platform itself, with visual feedback of the drawing available in real time to both participants. The narrator was able to say anything they wanted to the artist, providing further instruction as needed to complete the drawing. The artist's responses were limited by comparison. The artist was only able to communicate four pre-set phrases: 1) "Okay, what is next?; 2) "Could you please repeat that?" 3) "Can you clarify?" and 4) "My guess is X" (where they were allowed to freely type a word or phrase to complete the statement, Figure 27). The artist communicated each of these by clicking on-screen on the pre-set phrase. Similar to the narrator's instructions, the artist's response was said aloud in the voice they had chosen for themselves and was heard by both the artist and the narrator.

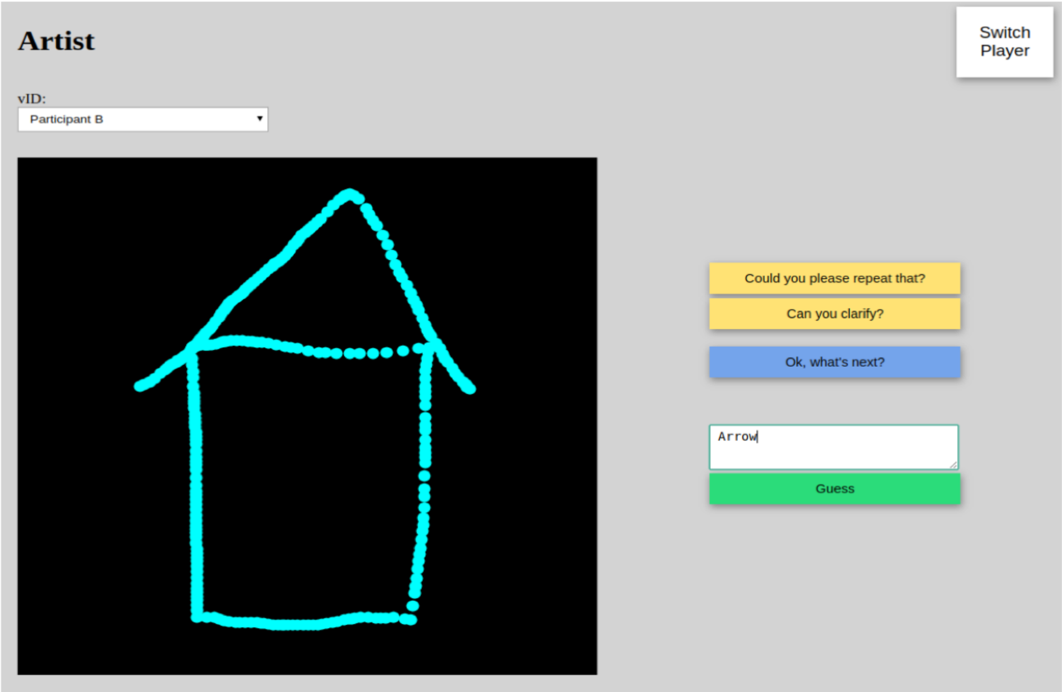

Figure 27. The participants' display when acting as the 'Artist' within the game, "Drawing conclusions".

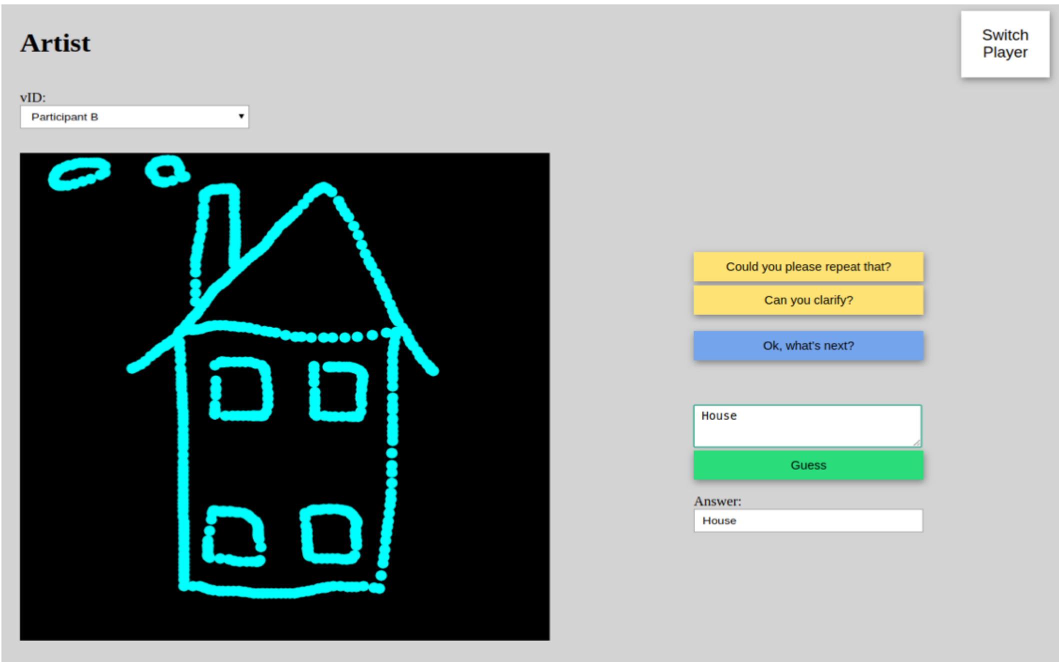

Figure 28. The participants' display when acting as the 'Artist' within the game, "Drawing conclusions". Here the artist has correctly guessed the image (e.g., House).

Limiting the artist's ability to communicate was done for two reasons: First, to minimise overlap between the two participants' voices during gameplay. Second, to better control for exposure to each of the participants' voices. For instance, when participant 1 was the narrator, both participants would have considerably greater exposure to that voice. However, after the artist had successfully guessed what the object was that they had been guided to draw (Figure 28), the two players switched roles. Thus, over the course of six rounds (each lasting ~ 5mins), each participant had acted as both the artist and the narrator three times. The game lasted for ~30 minutes for each participant.

Importantly, although each participant was told they would be paired with another participant, they were in fact paired with a researcher. This was done to ensure consistency in each participant's experience of the game. The researcher always began as the narrator in order to demonstrate how an object might be described (i.e. with an appropriate level of detail and with incremental instructions). Further, the researcher always described the same three pictures for the participant to draw, to ensure consistency in task difficulty and, where possible, to ensure that participants had a similar amount of exposure – and of similar content - to the researcher's voice. Given that the task was designed to provide an experience of the use of a self-voice in a social interaction, it was important that each participant felt it was a real human-human interaction rather than human-computer interaction. Thus, although the researcher's communications were largely pre-scripted, they were flexibly adapted ad-hoc to respond appropriately in each interaction. Within this task therefore, ecological validity was prioritised over controlling for exposure to each voice. However, the amount of exposure to each voice was recorded for subsequent analysis (see Exploratory Analysis). Each experimental participant therefore experienced having flexible use and extensive control over self-producing – and hearing – their newly chosen self-voice within the task. Further, each participant had experience of their fellow player's (the researcher's) chosen voice, which here constitutes a 'friend' voice.

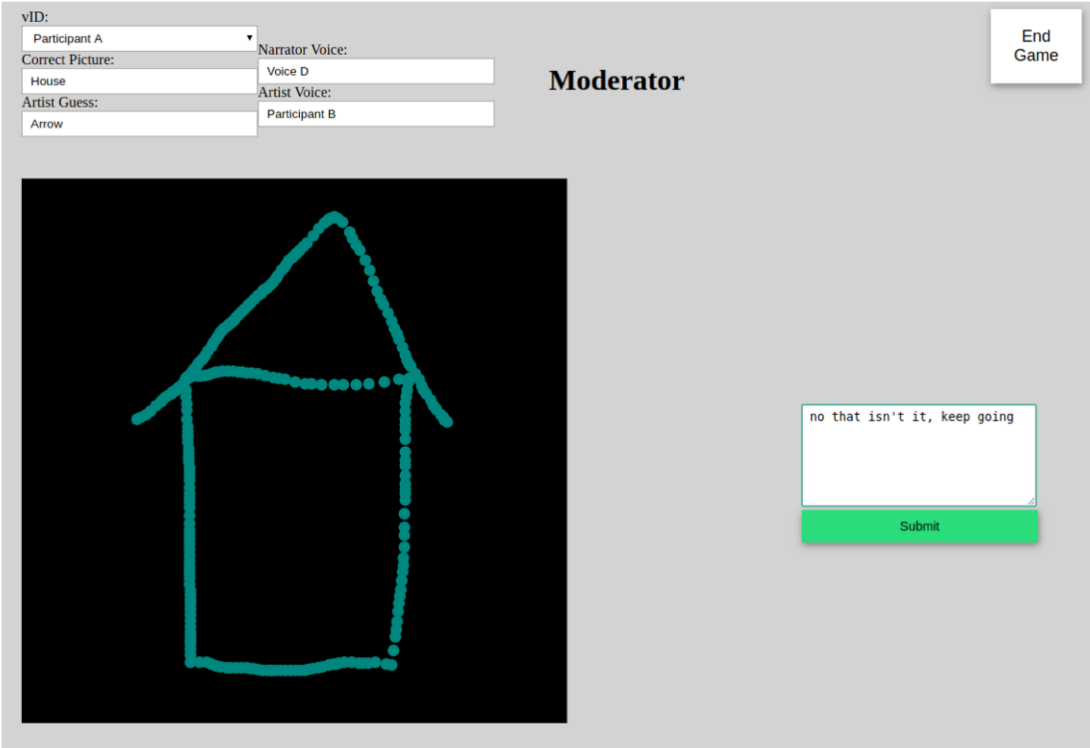

Figure 29. The Moderator's display within the game, "Drawing conclusions."

Finally, I included within the task a third voice, referred to as the moderator (Figure 29). It was through this moderator's voice that participants were given task instructions, told to switch roles at the end of each round, and guided to return at the end of the game to the test platform. This moderator voice was scripted but controlled live by a researcher. Communications from this voice were kept closer to the script as a more neutral third voice which could reasonably constitute a 'stranger' voice. Thus, within the task, participants experienced three social identities and their voices – the self, a friend, and an other – mirroring the three identities typically used within the perceptual matching paradigm reported in earlier studies within this thesis. Indeed, the perceptual matching paradigm was used here measure to self-bias and prioritisation over the self-voice, relative to the friend-voice (i.e., the fellow player) and an other-voice (i.e. the moderator) post game.

The voices assigned to the fellow player and to the moderator were counterbalanced across all participants. Both voices had an English accent, but one was female-sounding and the other male-sounding. After 6 complete rounds of the game, the game was ended and participants were directed back to the test platform, Gorilla.sc.

Task 3: Post-game questionnaire

After the drawing game, all experimental participants were asked about their experience of their self-voice, the friend-voice, and the other-voice. Specifically, participants were asked to:

1) To report whether they liked their self-voice i) more, ii) less, or iii) or the same after having used it;
2) How well they felt the self-voice had represented them on a scale of 1-7, where 1 = *not well at all* and 7= *very well;*
3) The degree to which they felt the other player was a friend on a scale of 1-7, where 1= *not at friendly all* and 7= *very friendly;*
4) The degree to which they felt the moderator was a friend on a scale of 1-7, where 1= *not at friendly all* and 7= *very friendly.*

Question 1 allows for the subsequent analysis of how personal preference may influence self-bias over a voice. Question 2 can provide insight on how self-bias is influenced by how accurately the new self-voice can represent the self. Questions 3 and 4 assess how the interaction itself may have influenced the social distance between self, friend and other. It is possible that participants experienced the fellow player as very friendly, which may reduce the social distance between self and friend and, in turn, reduce the difference in bias for each of the voices associated with those identities. These measures are analysed in Exploratory Analyses below (section 6.5.2).

Task 4: Perceptual Matching Task

All participants, both experimental and control, then completed a perceptual matching task to assess the bias accrued for each of the three voices by virtue of the identity they belong to. As in previous iterations of the task, all participants first completed a familiarisation phase. In this phase (<1 minute), participants were passively exposed to the three voices with an on-screen label showing the identity the voice was associated with (i.e., ‘YOU’, ‘FRIEND’, or ‘STRANGER’). The experimental participants were told that they would be reintroduced to the three voices they had already heard within the game; that here the voice of the other player was called a ‘friend’ voice and the voice of the moderator was called an ‘other’ voice.

The test phase then started with 12 practice trials. Thereafter, participants performed three blocks of 72 trials (216 trials in total) (for full procedural details, refer to Chapter 2, section 2.2.3.2). From this task, I measured reaction times and accuracy in recognising the voices associated with self, friend, and other when they were presented with the correct identity label (match trials) and with the incorrect identity label (mismatched labels).

Task 5: *Interval Estimation Task*

The final task was an interval estimation task, used to obtain an implicit measure of the sense of agency participants experienced over their self-voice relative to an other-voice. Within this task participants heard only two voices as possible outcomes: the self-voice and the other-voice. This was done to replicate previous versions of the task used in this thesis (see Chapter 3, section 3.2.1 for full details) and therefore did not include the friend-voice as a possible outcome. From this task, I measured participants’ estimations of the interval between a key-press or a cue-tone and an outcome, which was either the self-voice or the other-voice. As before, sense of agency is derived by the perceived compression of time between an action and its outcome relative to a passively heard tone and its outcome, as measured by interval estimates. Thereafter, differences in sense of agency over either the self-voice or the other-voice are derived by differences in estimates according to these two outcomes.

Overall, the five tasks comprising this study took participants 1 hour and 20 minutes, to complete, on average, with ~30 minutes of that time spent on the Drawing Game (Task 2) per participant.

6.2.3.2 Pre-processing

Perceptual Matching Task

To examine self-bias over a new self-voice I analysed three measures: reaction times, sensitivity, and accuracy.

All erroneous responses, as well as responses shorter than 200ms or longer than 1500ms were removed from the analysis in line with Sui et al.'s (2012) approach. This accounted for only 1.4% of trials. Whole data sets for four participants showing overall performance accuracy at chance (≤50% + 95% CI) were excluded and replaced as this indicated random responses and/or an inability to distinguish between the voice identities.

Interval Estimation Task

To measure the sense of agency, I assessed the interval estimations in both active and, separately, passive trials. To inform our exclusion criteria, I calculated a mean interval estimation per participant in each condition, e.g., Active trials with a 100ms interval were presented 10 times for each outcome (self-voice, other-voice) per participant and a mean estimate was calculated across these 10 trials. This mean estimation was used only to eliminate all estimates more than 2 standard deviations above or below each participant's mean (per condition) in line with previous studies (Engbert et al., 2008). 1.8% of trials were eliminated on these criteria.

Further, whole datasets of participants whose mean estimates did not increase monotonically with the presented interval duration (100ms, 400ms, 700ms) were also excluded and replaced as this demonstrated particularly difficulty with the task and/or a lack of engagement (Suzuki et al., 2019; Caspar et al., 2016). A further 5 participants were excluded on this criterion. Their data was replaced until a total of 88 (44 control, 44 experimental) participants had successfully completed the study.

6.2.3.3 Analysis

Perceptual Matching Task

The design consisted of one within-subject factor: 'voice identity' (*self* vs *friend* vs *other*) and one between-subjects factor: 'condition' (*experimental: self-produced voice* vs *control: self-owned voice*). Matched trials and mismatched trials were analysed separately.

Both reaction time data and sensitivity were assessed with linear mixed models (LMM) using lme4 (Bates et al., 2014) in the R environment (R Core Team, 2013). In both analyses, a model was run that included an interaction between *voice identity* and *condition* and a random intercept of *participant*, in match trials and, separately, mismatch trials.

E.g. lmer(reaction time ~ condition * voice + 1 | participant, REML = 'FALSE')

To assess accuracy in the perceptual matching task, I ran a binomial generalized linear mixed model (GLMM) that, again, included an interaction between *condition* and *voice identity* and a random intercept of 'participant', in match trials and, separately, mismatch trials.

E.g. glmer(accuracy ~ condition * voice + 1 | participant, family = "binomial")

Interval Estimation Task

Sense of agency was measured implicitly according to interval estimates within the temporal binding paradigm. The design consisted of three within-subject factors: 'voice identity' (*self* vs *other*), 'agency' (*active* vs *passive*), and 'interval' (*100ms* vs *400ms* vs *700ms*) and one between-subjects factor: 'condition' (*experimental: self-produced voice* vs *control: self-owned voice*).

Interval estimations were also assessed with linear mixed models. Specifically, I ran a model that included an interaction between *condition*, *agency* and *voice identity*, all possible two-way interactions and fixed effects of *condition, agency* and *voice identity.* Lastly, *interval* and *participant* were modelled as random effects.

E.g. lmer(interval estimation ~ condition * agency * voice identity + 1 | Interval + 1 | participant), REML = FALSE).

Statistical significance of the effects was established via likelihood ratio tests by dropping effects of interest from the full model which including the other fixed effects. For example, to establish whether an interaction was significant, the interaction was dropped from a model including the other possible interactions and the fixed effects. For all analyses reported here, post-hoc comparisons were conducted in emmeans (Lenth, 2016) and were adjusted for multiple comparisons via Bonferroni correction.
For LMMs, the models' estimates and associated confidence intervals for each effect are reported. The further away from 0 the estimates are, the bigger the effect. If the confidence intervals do not cross 0, the relevant effect is significant.

6.2.4 Results

6.2.4.1 Perceptual Matching Task

Descriptive statistics for reaction times, accuracy, and sensitivity scores (d') for experimental and control participants according to voice are given in Table 11.

***Table 11**. Mean RTs, accuracy, and sensitivity (d') in Experiment 10 (match trials).*

Condition	**Voice Identity**	**Mean RT (ms)**	**Accuracy**	***d'* ***
Self-produced voice	Self	524 (93)	0.96 (0.07)	2.91 (0.50)
	Friend	591 (116)	0.96 (0.07)	2.90 (0.54)
	Other	619 (121)	0.92 (0.09)	2.68 (0.63)
Self-owned voice	Self	546 (107)	0.96 (0.04)	2.92 (0.49)
	Friend	616 (116)	0.93 (0.10)	2.77 (0.74)
	Other	645 (133)	0.91 (0.09)	2.74 (0.60)

Note. RT = reaction time; Accuracy - proportion correct. Standard deviations appear within parenthesis. *Performance scores from match and mismatch trials are combined to provide d' scores.

Reaction Times

First, I assessed the reaction time data from the matched trials, plotted in Figure 30.

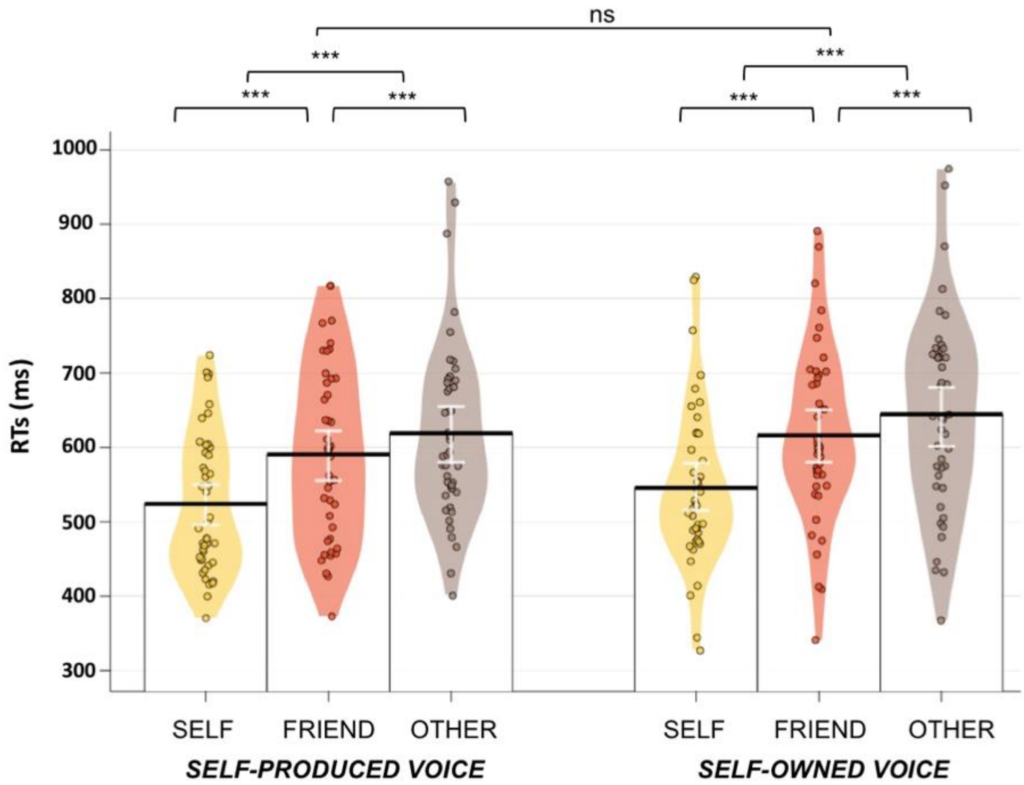

Figure 30. Experiment 10**:** Mean reaction times (RT) as a function of the voice identity (self, friend, or other) and condition (self-produced voice vs. self-owned voice). Graph models match trials only. The error bars indicate the SEs of the means. Coloured segments show smooth density curves for the full data distribution, while individual dots indicate mean RTs per participant. The top bar shows the interaction, horizontal bars show post-hoc comparisons. Asterisks denote significance within RTs as determined by likelihood ratio tests. *** p<.001

The interaction between *condition* (experimental, control) and *voice identity* (self, friend, other) was non-significant ($\chi2(2) = 0.194$, $p= .907$), suggesting that there was no detectable effect of having self-produced the voice on how much bias is afforded to it. There was, however, a significant main effect of *voice identity* ($\chi2(2) = 384.98$, $p<.001$) and post-hoc comparisons showed a significant self-prioritisation effect. Here, the voice chosen for the self was perceptually prioritised relative to either the friend-voice ($p<.001$) or the other-voice ($p<.001$) in both groups. Further, the friend-voice was perceptually prioritised relative to the other-voice ($p<.001$). The main effect of *condition,* however, was non-significant ($\chi2(1) = 1.236$, $p= .266$), indicating that overall reaction times did not differ significantly between experimental and control participants, according to whether they had self-produced their new voice, or not. Full model outputs are reproduced in Supplemental Table 20.

Overall, this shows that participants did attribute self-bias to a new self-voice such that it was prioritised in perception. Critically, though, the degree of bias afforded to the voice was not increased by having self-produced the voice in a social setting.

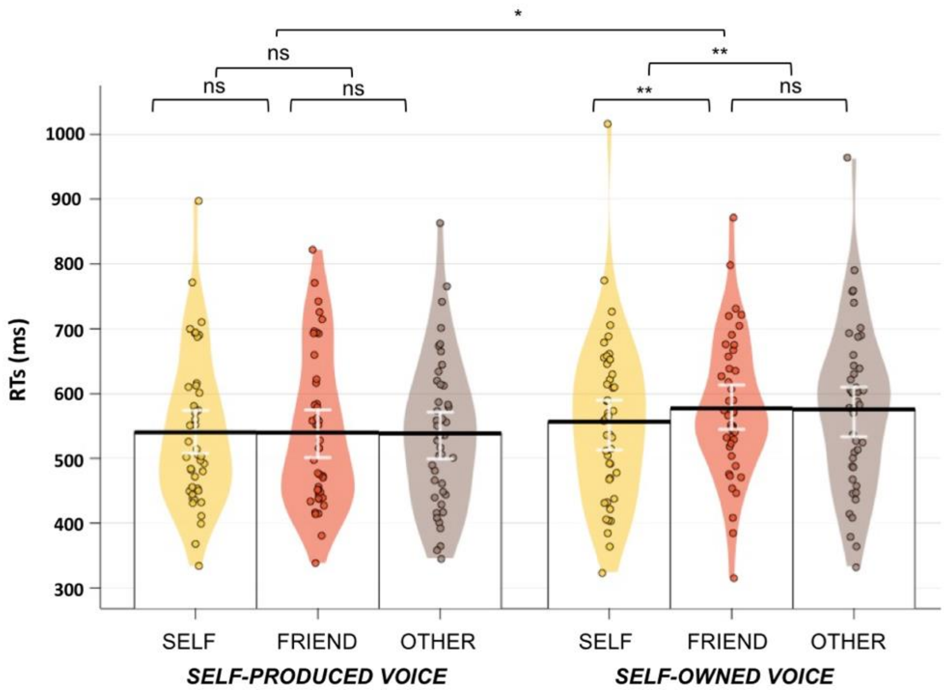

Figure 31. Experiment 10: Mean reaction times (RT) as a function of the voice identity (self, friend, or other) and condition (self-produced voice vs. self-owned voice). Graph models mismatch trials only. The error bars indicate the SEs of the means. Coloured segments show smooth density curves for the full data distribution, while individual dots indicate mean RTs per participant. The top bar shows the interaction, horizontal bars show post-hoc comparisons. Asterisks denote significance within RTs as determined by likelihood ratio tests. *** p<.001

In the mismatched trials, however, the interaction between *condition* (experimental, control) and *voice identity* (self, friend, other) was significant ($\chi2(2) = 6.44$, $p= .039$), see Figure 30). Post-hoc comparisons showed that this was because there was a main effect of *voice identity* in the control group but not the experimental group. Specifically, control participants showed a significant self-prioritisation effect such that the self-voice was judged more quickly than either the friend-voice (p=.002) or the other-voice (p=.004). This was not the case for the experimental participants, for whom reaction times did not differ between identities in the mismatch trials. This result is perhaps because experimental participants were more familiar with each of the three voices and better able to recognise each of them. The greater familiarity with all three identities may have increased the task difficulty in the mismatched trials and reduced the bias for the self-voice relative to the others.

Sensitivity

Full model outputs are reproduced in Supplemental Table 21. The interaction between *condition* and *voice identity* was non-significant ($\chi2(2)=1.88$, $p=.389$) This shows that self-producing the voice in a social context did not influence perceptual sensitivity to that voice in relation to the friend and other voices. However, there was a significant main effect of *voice identity* ($\chi2(2)=8.85$, $p=.012$); perceptual sensitivity was increased for the self-voice relative to the other-voice ($p=.009$) in both groups but did not differ significantly from the friend-voice ($p=.653$). Sensitivity also did not differ between the friend-voice and the other-voice ($p=.248$). Finally, the main effect of *condition* was non-significant ($\chi2(1)=.051$, $p=.821$). Overall, this shows that perceptual sensitivity was increased for the self-voice but this sensitivity was not influenced by self-producing the voice in a social setting.

Accuracy

Last, I ran a GLMM on trialwise accuracy, and full model outputs are reproduced in Supplemental Table 22. The interaction between *condition* and *voice identity* was non-significant ($\chi2(2)= 5.07$, $p=.079$). This showed that the effect of *voice identity* on perceptual accuracy did not significantly differ for experimental and control participants. Thus, greater familiarisation to the voices within the context of the Drawing Conclusions Game did not affect people's accuracy to the three voices overall. However, there was a significant main effect of *voice identity* such that, across groups, participants were more accurate at judging the self-voice relative to the friend-voice ($p=.009$) or the other-voice ($p<.001$) and more accurate at judging the friend-voice relative to the other-voice ($p<.001$). However, the main effect of *condition* was non-significant ($\chi2(2)= .059$, $p=.44$) showing that participants' bias toward the voices was not modulated by using the self-voice, nor by social exposure to the identities behind the friend-voice and the other-voice. Overall, there was a self-prioritisation effect in which accuracy for the self-voice is significantly greater, but that accuracy was not increased by self-producing the voice in a social setting.

In the mismatch trials, the interaction between *condition* and *voice identity* was non-significant ($\chi 2(2)=1.322$, $p=.516$), nor was the main effect of *voice identity* ($\chi 2(2)=4.173$, $p=.124$ and the main effect of *condition* ($\chi 2(1)= .201$, $p=.653$).

6.2.4.2 Interval Estimation Task

Sense of agency was then assessed by analysing the interval estimates measured in the temporal binding task. Mean interval estimations across all conditions are given in Table 12 and plotted in Figure 32a and 32b).

Table 12. *Mean interval estimates (ms) across conditions in Experiment 10.*

	ACTIVE TRIALS				**PASSIVE TRIALS**			
	EXPERIMENTAL		CONTROL		EXPERIMENTAL		CONTROL	
INTERVAL	**Self**	**Other**	**Self**	**Other**	**Self**	**Other**	**Self**	**Other**
100ms	169	202	140	146	246	264	236	252
400ms	327	341	318	333	416	443	424	437
700ms	485	516	512	558	613	627	643	671

The interaction between *condition, voice identity*, and *agency* was non-significant. Similarly, the two-way interaction between *condition* and *voice identity* was non-significant, which showed that the difference in estimates for intervals ending in the self-voice or the other-voice was not affected by whether participants has previously used the self-voice to represent themselves. Further, the interaction between *agency* and *voice identity* was also non-significant. This showed that the difference between interval estimates for the self-voice and the other-voice did not differ according to whether the interval occurred in an active trial or a passive trial. Similarly, the interaction between *agency* and condition was also non-significant, which indicates that participants who had agency in using their self-voice within the game demonstrated no greater sense of agency within the temporal binding task.

However, across all participants, the main effect of *agency* was significant ($\chi 2(1) = 922.05$, $p<.001$), showing that intervals were perceived to be shorter in the active trials compared to the passive trials. It is this difference between active and passive trials that is typically indicative of a temporal binding effect and this measure is used to derive peoples' sense of agency over outcomes that are self-generated relative to passive trials. Finally, then, there was a significant effect of *voice identity* ($\chi 2(1) = 42.61$, $p<.001$), with post-hoc comparisons showing that participants – in both experimental and control conditions, across both active and passive trials – estimated intervals that terminated in the self-voice to be significantly shorter in duration that those terminating in the other-voice ($p<.001$). While this clearly shows that interval estimates were sensitive to whether an outcome was self-owned or other-owned, this status of ownership affected both active and passive trials as in Experiments 4, 5, and 7. If it is the participants' sense of agency with which self-ownership interacts, this effect should be present only in the active trials. However, self-ownership also affected the passive trials here, which suggests that it may be influencing interval estimates by affecting temporal perception more broadly and not necessarily people's sense of agency.

Overall, the results suggest that the experience of producing a new self-voice did not influence participants' sense of agency over that voice within this task. This tallies with the results from the perceptual matching task. Taken together, the planned analyses in this study suggest that the use of a new self-voice – both as a means of self-representation and of communication – did not influence the amount of self-bias afforded to it, nor the sense of agency over it.

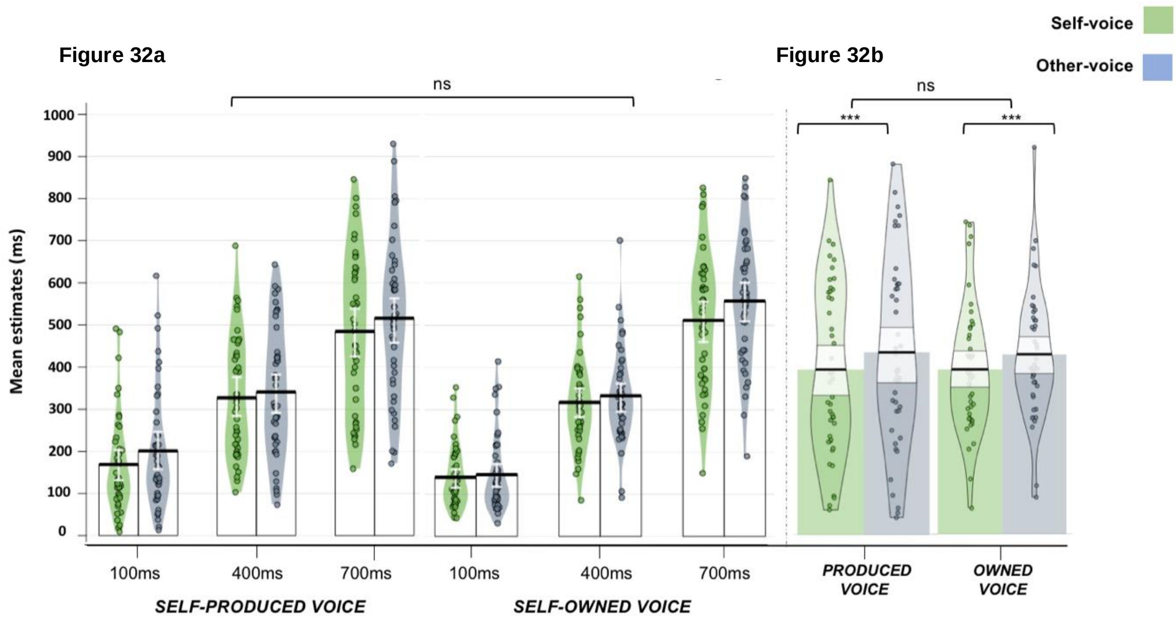

Figure 32a. Experiment 10**:** Mean interval estimates (ms) according to the outcome (self-voice or other-voice), and the condition (self-produced vs. self-owned). Intervals are plotted according to the true interval duration (100ms, 400ms, and 700ms) and for active trials only. The error bars indicate the SEs of the means. **Figure 32b.** Experiment 10**:** Mean interval estimates (ms) according to the outcome (self-voice or other-voice), and the condition (self-produced vs. self-owned). Estimates are collapsed across interval duration. Inference bands indicate CIs. Coloured segments show smooth density curves for the full data distribution, while individual dots indicate mean estimates per participant. Top bars show interactions, horizontal bars show post-hoc comparisons with asterisks denoting significance within estimates as determined by likelihood ratio tests. *** p<.001

6.2.5 Exploratory Analyses

Considering the current result, it was important to further explore how individual participants' experience of using the self-voice may have influenced their self-bias and agency. Given the emphasis on ecological validity and, hence, the complex social task, there may be underlying individual differences at a group level that obscured the presence of the effects of social use of the voice in some participants. As outlined earlier (see section 6.2.3.1: Procedure), participants were asked whether they liked their new self-voice more, the same, or less after using it to represent themselves within the task, and also whether their chosen voice matched their accent. Further, participants were asked to report on a Likert scale (1-7, whereby 1 denoted '*Not at all'* and 7 denoted '*Very'*) how representative they felt their chosen voice had been. Finally, on the same scale, participants were asked how friendly they perceived the other player and the moderator to be. Responses to the questionnaire by experimental participants (i.e., those who used their self-voice within the game) are reported below in Tables 13a, b, and c.

Table 13a. *Questionnaire responses: Participant responses according to their attitude towards the self-voice in Experiment 10.*

Like the self-voice:	No. of participants
Less	14
Same	12
More	18

Table 13b. *Questionnaire responses: Participant responses according to whether their new self-voice matched their accent or not in Experiment 10.*

Matched	No. of participants
Yes	14
No	30

Table 13c. *Questionnaire responses: Participant responses according to self-voice representativeness and perceived friendliness of players in Experiment 10.*

Likert Score	**1**	**2**	**3**	**4**	**5**	**6**	**7**
How well the voice represented the 'self'	2	3	3	5	13	6	12
Perceived friendliness of the other player	0	0	0	1	1	14	28
Perceived friendliness of the moderator	0	6	3	14	9	4	8

It is important to note that some of the sub-grouped samples arising from this questionnaire were small and imbalanced. Thus, in some instances it was not feasible to use the questionnaire data for further analyses. The vast majority of participants rated their fellow player as very friendly (i.e., a Likert score of 6 or 7 out of 7) and only two participants rated them to be less friendly. This meant that further testing of the role of perceived friendliness of the "friend" was not viable. This measure could have provided insight into how another may be attributed more or less bias according to their perceived social distance from self, which may have been reduced by the other being friendlier and co-operative with the self. Thus, future work will require larger sample sizes which may yield larger sub-groups according to perception of the ‘friend’ and ‘other’.

However, I ran two further analyses. First, exploring whether the amount of exposure participants had to each of the three voices had an effect either on how participants attributed bias to them or experienced a sense of agency over them. Second, on whether participants attitude toward their self-voice i.e., whether they liked it more, less, or the same as before they had used it, affected how they prioritised it or experienced agency over it.

6.2.5.1 Does the extent of experience using the self-voice influence self-prioritisation or agency?

There was substantial variation in participants’ individual usage of the self-voice (see Figure 33) and therefore in their exposure to – and familiarity with – the self-voice relative to the other voices.

I therefore wanted to assess a) whether participants who used their self-voice more may have attributed greater bias to it and; b) whether greater familiarity with the friend-voice or other-voice (according to how much each voice was heard) subsequently affected bias towards each of them. The amount each voice was used was recorded by logging all of the typed text produced in the game. For each participant, the number of words recorded for the self-voice equates to the number of words they, themselves, typed and produced. The number of words recorded for the friend-voice and other-voice relates to the number of words that participant heard aloud for each of those voices, and thus provides a measure of exposure and familiarity. Notably, exposure to the self-voice was relatively less on average than exposure to either the friend-voice or other-voice.

Figure 33. Experiment 10**:** Mean number of words heard per voice identity (self, friend, or other) within the game, “Drawing Conclusions”. Coloured segments show smooth density curves for the full data distribution, while individual dots indicate mean RTs per participant.

An average reaction time to the self-voice, friend-voice and other-voice was calculated per participant from the perceptual matching task. Further, a measure of exposure to each voice was calculated according to the number of words produced and/or heard aloud in that voice, per participant within the game. The relation between these two measures was initially assessed with three Spearman’s rank-based correlation tests. First, I assessed whether the amount of exposure to the self-voice was correlated with reaction times to that self-voice. The results showed that the measures were not significantly correlated ($r(42)= .030$, $p = .845$). This shows that participants’ reaction times to the self-voice – and therefore the bias they attributed to it – were not influenced by how extensively they produced the self-voice in a social setting.

Second, I assessed whether the amount of exposure to the friend-voice was correlated with the reaction times to the friend-voice but, again, there was no significant correlation between the two measures ($r(42)= -.014$, $p=.929$). Finally, I assessed whether the amount of exposure to the other-voice was correlated with the reaction times to the other-voice. This was also non-significant (r(42)= -.039, p=.798). Overall, these tests show that the amount of exposure participants had to each of the three voices did not directly affect reaction times to those voices in the subsequent perceptual matching paradigm. This suggests that participants' prioritisation of each voice was independent of how much exposure they had had to them.

However, it is possible that the voices were differentially prioritised *relative* to each other, according to how much exposure was had to each voice. To assess this, I first calculated the degree to which the self-voice was prioritised relative to the friend-voice. Following Sui & Humphreys (2017), this was done by taking the difference in RTs for the self-voice versus the friend-voice, dividing it by the sum of RTs across the two conditions, and finally multiplying the result by 100 to achieve a percentage score. This allowed a measure of bias for the self-voice relative to the friend-voice to be generated. A measure of exposure to the self-voice relative to the friend-voice was similarly calculated by taking the difference in the number of words for the self-voice relative to the friend-voice, dividing it by the sum of words across both voices and then multiplying the result by 100. These two measures were then assessed in a Spearman's rank-based correlation test and the two measures were non-significantly correlated (r(42)= .039, p=.796). This suggests that the degree to which participants prioritised the self-voice over the friend-voice was not influenced by the different levels of exposure to the self-voice relative to the friend-voice within the game.

Thereafter, this was process was repeated for the voice identities. Specifically, when the difference in exposure to the self-voice relative to the other-voice was analysed against the difference in bias for the self-voice relative to the other-voice, the spearman rank-based test was non-significant again (r(42)= .088, p=.567), as was the difference between the friend-voice and the other-voice (r(42)= .041, p=.789). This shows that the different amounts of exposure to each of the voices did not affect the way in which those voices were perceptually prioritised relative to one another.

It is interesting then that each participant was actively involved in the study for ~1 hour and 20 minutes hours and, particularly, in the Drawing Conclusions game for ~30 minutes. Within this time, the average exposure to their chosen self-voice was only 213 (sd = 78) words across an average of 31 (sd = 8.2) utterances. Given that the speech rate of text-to-speech voices is around 200 words per minute, this shows that participants actually had relatively little exposure to their voice. Holmes, To, & Johnsrude (2021) recently conducted a study into the threshold at which voices became familiar enough to confer an advantage to speech intelligibility and voice recognition. Within this study, 78 sentences roughly corresponded to 10 minutes of exposure to a new voice. This exposure was sufficient for a voice to become more intelligible than a novel voice when in the presence of a competing speaker. Moreover, after 60 minutes of exposure, this intelligibility benefit increased 10-15% which is on par with the benefit previously found for personally familiar voices belonging to family and close friends. Here then, although participants were engaged in the task for 30 minutes, this was commensurate to only a couple of minutes of exposure to each voice, which is well below the minimum threshold in Holmes, To, & Johnsrude (2021). Thus, it is possible that the voices were not produced or heard enough for their familiarity to confer a benefit. This is supported by the fact that there was enhanced perceptual prioritisation for the highly familiar self-voice relative to a new self-owned voice in Chapter 4. That enhanced prioritisation is not evident here; there was a similar pattern of prioritisation for all three voices by experimental participants and controls, for whom they were novel. This suggests that the voices have not become familiar enough for familiarity to confer a benefit.

Overall, this corroborates the result that self-producing the voice does not influence subsequent bias towards it. However, within this task, participants on average had greatest exposure to the other-voice and least exposure to the self-voice. Still, there was a significant self-prioritisation effect such that participants recognised the self-voice more quickly and accurately than the other voices. This shows the robustness of this self-bias effect; that, by a voice becoming self-associated through ownership, it becomes self-relevant enough to be perceptually prioritised above voices that might be perceptually more familiar.

It was also important to ask whether the extent to which the self-voice was produced by participants influenced their sense of agency over it. Within the task, participants were able to have agency and control over using their new self-voice, which could be produced by typing in text and pressing submit. This action of pressing keys to generate the voice was then replicated in the temporal binding task; participants made a keypress and heard either the self-voice or the other-voice as an outcome of that action. Thus, the method of self-production was comparable across the two tasks. Intuitively then, greater experience of having agency in producing the self-voice via these means may have increased the sense of agency participants had over that voice thereafter. Given that there was substantial variation in the extent to which participants used their self-voice, I here analysed whether the extent of self-production was correlated to the later sense of agency over that voice.

To analyse this, a measure of agency for each voice was calculated by taking the difference in estimates for trials in the active condition relative to the passive condition, for both the self-voice and the other-voice at each delay, per participant. First then, the measure of agency for the self-voice was submitted to a Spearman's rank-based correlation alongside the number of words self-produced by each participant. The analysis showed that the two measures were non-significantly correlated ($r(42)=.0003$, $p=.997$). This suggests that the extent to which participants had prior experience of using the self-voice did not affect the sense of agency they experienced over the voice in the interval estimation task.

The agency measure for the other-voice was submitted to a Spearman's rank-based correlation alongside the number of words heard in the other's voice and the two measures were weakly, negatively, correlated ($r(42)=-.218$, $p=.015$). The analysis showed that the more words participants were exposed to in the other-voice, the less agency they had over that voice. Interestingly, participants who were more exposed to the other-voice will have had greater interaction with that other-voice. The other-voice was representative of a moderator and so it is probable that this greater interaction was underpinned by the moderator's script needing to be adapted to respond to idiosyncratic issues within the task. This necessitated more human-like responses as opposed to the more neutral base script. Thus, in cases in which participants had relatively greater exposure to the other-voice, it is likely that they also experienced more human-like agency and responses from that voice. As a result, people may have experienced less agency over the other's voice as it became more associated with another agent (i.e., a human-like moderator) as opposed to a computer. Previous studies have shown that when in the presence of other agents, people's sense of agency can be reduced (Beyer, Sidarus, Bonicalzi & Haggard, 2016; Ciardo, Beyer, De Tommaso & Wykowska, 2020). This is because the responsibility for the outcome can be attributed to that human agent and so the participants' own sense of agency is reduced as is it diffused across agents.

Overall, these analyses suggest that using the self-voice did not affect the degree to which participants prioritised it perceptually or felt a sense of agency over it. However, it is possible that familiarity with other voice identities may affect the extent to which they are perceived as being underpinned by an agent and may, therefore, diminish temporal binding over other-associated outcomes.

6.2.5.2 Does (dis)liking the new voice influence bias or agency?

Within the task, any shifts in attitude regarding the new voice were recorded after participants' experience of using it. The modal response (41% of participants) was that participants liked the voice more as a result of having used it to interact socially. A further 27% of participants reported that there had been no change; they liked the voice as much having used it as when they first chose it. Lastly, just under a third of participants (32%) reported they liked the voice less after using it within the drawing game than they did when they had first chosen it.

It is possible that people who liked their new self-voice were more biased towards it relative to people who did not like their self-voice as much. If they liked it less, it may have been rejected as a self-voice, which may also have reduced their bias towards it. I therefore assessed whether the self-prioritisation effect was modulated according to how much the self-voice was liked.

To analyse this, participants were categorised into three groups based on their attitude towards the voice (liking: less, the same, more). A linear mixed model was then run on reaction times with *attitude* (less, the same, more) modelled as a predictor alongside *voice identity* (self, friend, other):

E.g. lmer(reaction time ~ voice * attitude + 1 | participant, REML = 'FALSE')

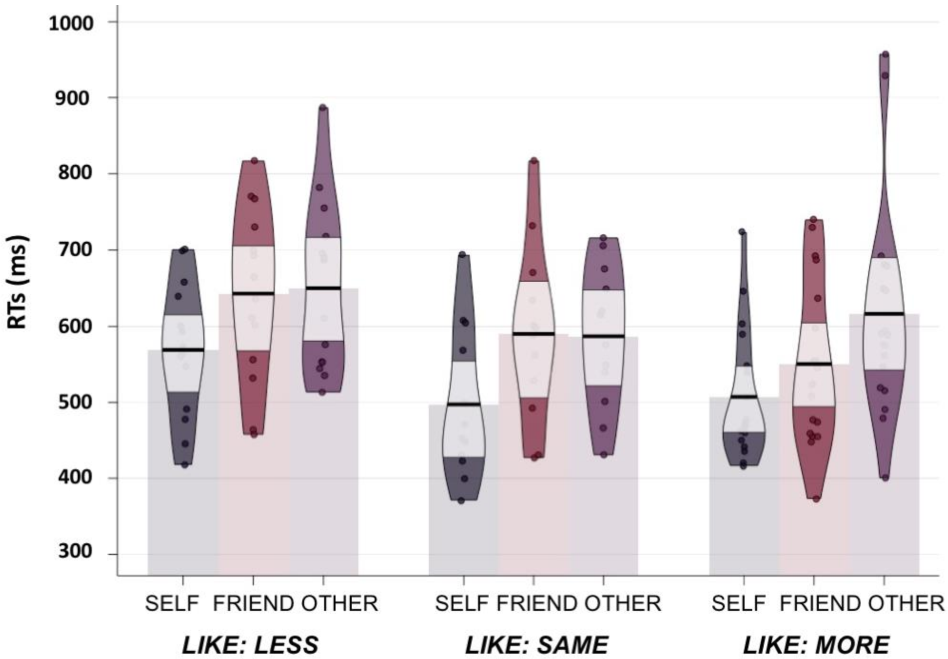

Figure 34. Experiment 10**:** Mean reaction times (RT) as a function of the voice identity (self, friend, or other) and attitude towards the self-voice after using it to represent the self (like it: less, the same, more) **Graph models match trials only**. Coloured segments show smooth density curves for the full data distribution, while individual dots indicate mean RTs per participant.

The reaction time data are plotted in Figure 34. The interaction between *voice identity* and *attitude* was significant ($\chi 2(4) = 21.16$, $p<.001$). The interaction arose because only in the group of participants who liked their new self-voice more, did their reaction times differ between the friend-voice and the other-voice ($p<.001$). In this group, the friend-voice was prioritised similarly to the self-voice and so the perceptual distance between the friend and other was increased. Thus, when participants liked their own self-voice more, they also showed a bias towards the friend's voice. This result may be underpinned by the success of the social interaction between the self and the friend. If participants found their new self-voice enabled a particularly positive interaction with another participant, they may report liking their own voice more. Further, it is likely that the social distance between the two players will have been reduced; in being a co-operative partner, the friend voice may be deemed similar to self and a bias afforded them as a result. For participants who reported liking their own voice less after use, or only as much as when chosen, the friend is relatively deprioritised and processed similar to the other-voice (ps>.05).

Interestingly, participants' prioritisation for their own self-voice was unaffected their attitude towards the voice. The post-hoc comparisons showed that participants' who liked their self-voice more did not prioritise it to any greater degree than participants who liked their voice less (p=.590) or liked it the same as before they had used it (p=.819). Similarly, participants who liked their voice less did not prioritise it to any different degree than participants who liked it the same as before using it (p=.057). This suggests that the bias for the self-voice is relatively automatic and may be more robust against influences that could lead to deprioritisation of 'self'.

Overall, this shows that people's attitude towards their new self-voice did not influence the extent to which they attributed bias towards it and perceptually prioritised it. This suggests that liking the voice more did not increase how self-relevant it was perceived to be. There was, however, a relation between how much the self-voice was liked and how much the friend-voice was prioritised. While it seems unlikely that these are directly related, it is probable that they are both symptomatic of the experience of social interaction more widely. Individual differences in their experiences of the game may directly affect people's perception of both the self, the other, and the relation between the two, which is evident within the behavioural measures here.

It is also possible that there could be a relationship between participants' sense of agency and how much they liked their new self-voice. I therefore analysed interval estimates from the temporal binding task with a linear mixed model as below:

E.g. lmer(estimates ~ agency * voice identity * attitude + 1 | participant + 1 | interval, REML = 'FALSE')

Here, the three-way interaction between *voice identity*, *agency*, and *attitude* was non-significant. Further, all two-way interactions including the factor of *attitude*, and the main effect of *attitude* were all non-significant. This shows that participants' sense of agency over all of the outcomes they generated, both the self-voice and the other-voice, was not significantly modulated by the extent to which they liked their new-voice. Unlike participant's perceptual prioritisation then, their sense of agency was comparatively unaffected by their attitudes towards their new self-voice. The null effect makes sense; people can like someone else's voice without feeling a sense of agency over it. However, this analysis also shows that attitudes towards the self-voice did not influence how readily it was accepted as an outcome to self-generated action.

Overall, people's individual attitudes towards their new self-voice affected how that voice was prioritised perceptually but not their sense of agency over it.

6.3 Discussion

Within the current study, I have demonstrated that the use of a new self-owned voice in a social interaction did not modulate the self-bias attributed to it, nor the sense of agency participants had over it. Given that it is through using the self-voice that people are able to represent themselves, build social relationships, and achieve communicative goals, this may be surprising. However, this finding does support the previous results presented in Payne et al. (2020) and furthers our understanding of how self-bias is attributed.

6.3.1 Perceptual prioritisation of self-voices

Most prominently, the results indicate that people are biased towards their self-voice because it is something they own, and not necessarily because it provides a relevant means of self-expression. Indeed, mere ownership over a voice was sufficient for self-bias to be attributed to it and for it to become processed as a self-relevant stimulus. The degree of perceptual prioritisation afforded to the self-voice was similar for participants regardless of the fact that the voice had been self-produced, how familiar the voice became or how much it was liked.

However, it is important to consider that the self-prioritisation effect has been modulated in previous studies. For instance, prioritisation for a self-associated stimulus was greater if it was associated with a "good" aspect of the self, relative to a "bad" aspect of the self (Hu, Lan, Macrae and Sui, 2020). Specifically, participants were asked to associate a stimulus with the morally good parts of themselves and another stimulus with the immoral parts of themselves. At test, participants showed faster and more accurate reactions to the "good" self. Further, in a recent study by Wang, Qi, Li and Jia (2021), perceptual prioritisation was increased for a stimulus that was associated with both the self and with a monetary reward, relative to a stimulus that was only associated to the self *or* a reward. The authors suggest that the condition of "double salience" (i.e. self-relevant and reward-associated) drives stronger prioritisation.

In both studies though, each participant was presented with the two associations as competing stimuli. Such studies examine the degree of self-bias attributed to different stimuli *within* participants and, indeed, find that some self-stimuli are more or less prioritised than other self-stimuli according to their features. In the current study, however, I examined the degree of self-bias attributed to a self-produced voice by experimental participants in comparison to the degree of self-bias attributed to a self-owned voice by control participants. The results showed that participants in each group attributed a similar level of bias to their new self-voice, but they only ever experienced one new self-associated voice. If, however, the self-produced voice and a self-owned voice were contrasted directly within participants, it may be that one is prioritised more than the other.

Indeed, this is evident in a study by Enock, Hewstone, Lockwood and Sui (2020) in which a self-prioritisation effect was evident for a self-associated stimulus and, separately, for a stimulus associated with a self-relevant team (i.e., a bias for an in-group stimulus). However, when the two stimuli were competing, there was significantly reduced bias for the team-associated stimulus so that only the self-stimulus was prioritised. While the current study demonstrated that people do not afford greater self-bias to a self-voice when it is self-produced, it is still possible that this voice would be relatively more prioritised than another self-voice that was only owned. This within-subject comparison could form the basis of future self-bias studies.

The current study also provides insights into how our wider perceptual processing is affected by self-association. Previous studies have shown that self-produced outcomes tend to be more highly valued and more positively regarded. This is because such outcomes can reflect an individualised and competent self. Here then, participants were provided with a new means of self-expression that *could* signal their competence through its use, but that does not necessarily mean it did. A third of participants reported liking their voice less after having self-produced it, which suggests that people did not automatically regard the voice more positively than other voices purely because it had been self-produced. Critically though, the self-voice was similarly prioritised by all participants regardless of how positively they regarded it.

Finally, in self-producing the voice, it should have become more familiar. Experiment 6 presented earlier within this thesis (Chapter 4) demonstrated that the true self-voice was prioritised relative to a voice that was only self-owned. This result was assumed to be because the true self-voice was more familiar and/or because it was self-produced. For the experimental participants within the current study, the new self-voice became a voice that was self-produced and, through its use, more familiar than it was for control participants who only owned it. Yet self-producing the voice and its increased familiarity did not increase how highly it was prioritised here; both experimental and control participants processed it similarly. This is at odds with the earlier experiment featuring the true self-voice. Given that participants had the least exposure to the new self-voice relative to the friend-voice or other-voice, it may that the level of exposure was not sufficient for familiarisation to that voice to imbue greater bias. Indeed, the true self-voice is a highly, personally, familiar voice and it is possible that exposure to the new self-voice through its use was not equivalent.

The fact that the greater bias was not afforded to a new self-voice because it could be self-produced may also be due to the way the voice was produced. All the participants who completed the study have a biological self-voice which they retain flexibly control and agency over. Therefore, controlling a new self-voice via text-to-speech technology may never be a comparable means of self-producing speech. If this study was run with users of AAC devices who required an alternative voice that they could control, then its self-production via a technological interface may lead to increased prioritisation relative to a voice that could not be produced.

6.3.2 Sense of agency over self-voices

This study also extended the examination of sense of agency over a new self-voice. Despite participants having prior experience of agency and control over self-producing the voice, this agency was not reflected in the temporal binding results. Rather, the degree of temporal compression for intervals terminating in the self-voice (relative to the other-voice) was the same for those who had produced the self-voice as for those who had only owned it. Thus, prior experience of the self-voice as a reliable outcome of self-produced action did not increase this agency.

Relatedly, it must be acknowledged again that the effect of *voice identity* – the difference in estimates for intervals terminating in the self-voice, relative to the other-voice – was present in both active and passive trials and across experimental and control participants. This shows that the effect of *voice identity* may have modulated temporal judgements outside of conditions of agency. The fact that participants' interval estimates were significantly reduced for the self-voice may therefore not be because self-ownership increases sense of agency *per se*, but rather because self-relevance of a sensory stimulus interacts with perceptual processing more broadly. Further studies will be required to determine this.

Overall, this study sheds lights on how we process a new self-voice, not only as a voice we have ownership over but as a voice we can use to represent the self. I showed that ownership over the voice is sufficient to bias perception towards it as a self-relevant stimulus. In future studies it will be interesting to increase the amount of use participants have over their voice and determine whether more prolonged production increases agency. Moreover, it will be important to examine bias and perceptual processing of a new voice by people whose current method of self-production similarly relies on text-to-speech-technology. The relevance of a new voice and its importance to the self as a possible means of self-representation may have been limited in participants recruited within this experiment. Further studies may therefore examine whether the relevance of a new voice as an alternative means of self-expression is modulated not by the voice itself, but by the person who uses it.

7 General Discussion

In this thesis I have explored the question of what it is that makes a voice ‘mine’ and provided the first behavioural investigation into the possibility of incorporating a new voice into the self-concept.

I have examined the influence of giving people ownership of a new voice, personal choice in selecting it, and agency in using it, on the degree to which people will accept this unfamiliar voice as being ‘self’. I have assessed the extent to which a new voice has become accepted as ‘self’ according to three main lines of enquiry. First, by examining whether it is prioritised as a self-relevant stimulus in perception; second, by examining whether people experience a greater sense of agency over it, and; third, by assessing whether there is a memory advantage for information expressed in the new voice.

Below I discuss the overarching conclusions that can be drawn from the ten behavioural studies that comprise this thesis, and their implications. Thereafter, I discuss the limitations of this work and suggest avenues of future research.

7.1 The Role of Ownership

Common to all the studies within this thesis is that participants gained ownership of a new self-voice. Previous studies have shown that the self can be extended through ownership, such that self-owned stimuli are processed preferentially as self-relevant. This effect of ownership is shown extensively here.

In Experiment 1, I demonstrated that an unfamiliar voice, which is the inherent biological property of another, can be made relevant to the self through ownership. By accepting that a new voice temporarily, and indeed arbitrarily, belongs to the ‘self’, it is prioritised in perception as a signal that is more relevant than the voices of others. This suggests that what is ‘self’ has been extended to encompass this new voice, which would previously have been processed as signalling an ‘other’. This aligns with previous literature which suggests that the boundaries of self are extended by what we own (Belk, 1988; Heersmink, 2020; Mittal, 2006). This has similarly been shown for faces; people who are told another's face temporarily belongs to them then prioritise it as self-relevant (Payne et al., 2017). Yet, although both the self-face and self-voice are primary means of self-representation, the voice can only exist through self-generated action. Thus, Experiment 1 provides the first demonstration of the extent to which the self can be extended through ownership to incorporate new other-related stimuli into the self. Indeed, through ownership, another's voice can become 'mine'.

This result is further supported by all the experiments within this thesis that tested prioritisation of a new self-voice. In every experiment, when an unfamiliar voice became self-owned, it was perceptually prioritised. There were consistently faster reaction times and greater accuracy for a new self-owned voice relative to the voices assigned to other social identities, as shown within a perceptual matching paradigm. This effect of ownership on making a voice self-relevant enough to be prioritised was robust; the new voice was prioritised as 'self' regardless of whether it matched participants' gender-identity (Experiment 2); whether the voice was chosen or arbitrarily assigned (Experiment 3) and, moreover, regardless of how much people liked the voice or whether it could be used as a means of self-expression or not (Experiment 10).

The fact that, through ownership, the new self-voice has been processed as self-relevant is corroborated by Experiments 6 and 7. In these experiments, participants heard their true self-voice within the task either as belonging to the self (Experiment 6) or as belonging to an 'other' (Experiment 7). When their true self-voice was assigned as being self-owned, it was prioritised over both the friend-voice and other-voice. Importantly, it was prioritised over these other voice identities relatively more than the degree to which a new self-owned voice was prioritised over these other voice identities. While this initially suggested that that the true self-voice may be deemed more self-relevant than a new self-voice, the bias for the true self-voice was diminished when ownership of it was transferred to an 'other'. This demonstrated the role of ownership in guiding our perception of what is self or other, and therefore, in biasing our perception towards what is self-relevant or not.

Importantly, I demonstrated that a small bias was evident for a new self-owned voice (relative to a friend's voice) even when that new voice was competing with participants' true self-voice. Thus, even in the presence of what is truly self-relevant, a new voice can be attributed a level of self-bias and prioritised as self-relevant because it is self-owned.

The role of ownership did not, however, affect memory, as there was no evidence of a memory advantage for self-owned items relative to other-owned items. In Experiments 8 and 9, I asked whether people demonstrated better memory for items said aloud in a new self-owned voice relative to an other-owned voice. However, there was no evidence to support this hypothesis, which suggested that the bias that is evident in perception did not extend to memory. However, this null result may have been related to the task design. The finding that there is better memory for self-owned items has been replicated extensively in the previous literature and yet, here, the same effect was not found. The status of items as being either self-owned or other-owned was signalled by them being heard in the self-voice or the other's voice. This assignment of ownership – of flagging what was, or was not, self-relevant – may have been too implicit. Further, Turk, et al. (2013) posit that the encoding of self-relevant items requires increased attentional resources. Thus, it is likely that within the task design the self-relevance of potentially self-owned items was not salient enough and/or participant engagement with the task not sufficient for effective self-referential encoding of so many new stimuli.

However, the effect of having ownership of a new voice did extend to influencing participants' perception of time and sense of agency. Sense of agency is the feeling of authorship and/or belief that "I caused that". Within this thesis, sense of agency was derived from a temporal binding effect as shown by interval estimation tasks. The results consistently revealed that people perceived the duration between an action and its outcome as being shorter if the outcome was the self-owned voice, relative to the voice of another. According to previous literature, this indicates that participants have a greater sense of agency over the self-voice.

On these results alone, it could be interpreted that ownership of a new voice is sufficient for people to experience a sense of agency over it. Indeed, an increased sense of agency was evident in instances in which ownership was afforded by being arbitrarily assigned, as well when voices were personally chosen. Further, when a participant's true self-voice was also presented in-task but as an other-owned stimulus, still there was a relatively greater sense of agency over the new self-voice. In reality, the true self-voice is also self-owned, so the fact that there was greater agency over the *new* self-voice within the tasks shows that temporal binding is sensitive to the immediate context and to what is, in that instance, most self-relevant.

However, a caveat to these results is that it is unclear whether the difference between interval estimates for durations preceding a self-voice outcome and an other-voice outcome are, truly, indicative of a difference in agency. Whether a voice was self-owned or other-owned frequently affected participants' temporal judgements similarly in active and passive trials across the experiments presented within this thesis. This body of work may be the first to show that factors which give rise to different degrees of binding in active trials – which might alone be interpreted as indicative of sense of agency – may also affect binding in a condition of non-agency. Indeed, previous temporal binding studies which do not use a passive condition and instead interpret differences *within* active trials only may need to elucidate whether these factors exclusively modulate temporal judgements under conditions of agency.

This research additionally demonstrates that self-generated actions can still have different degrees of self-relevance and elicit a different degree of temporal binding. A self-owned voice resulting from self-generated action affected interval estimates differently from an other-owned voice resulting from self-generated action. Within previous temporal binding paradigms, outcomes of self-generated action (i.e., key-press) are typically assumed to be self-relevant because they have been self-produced. This thesis highlights the importance of clarifying how the self-relevance of those outcomes interacts with participants' temporal perception. Indeed, it is possible that temporal perception, more generally, is biased by self-relevance and this needs further exploration.

Together though, these results suggest that the knowledge that the self is related to a stimulus through ownership of it – the knowledge that it is 'mine' – quickly and pervasively shapes our experience, and predominantly our perception of that stimulus. From these studies it is clear that the role of ownership is that it is a means through which new stimuli can be made self-relevant. According to Chiu, Ho, Tollenaar (2021), the self-concept comprises 'internal representations that shape perceptions of how the self is related to one's surroundings and to other people,' (pg. 1). Thus, through ownership, people can flexibly shift how they are related to external things, such that owning an item is to be related to that item.

7.2 The Role of Choice

In Experiments 1 and 2 participants were randomly assigned ownership of a new self-voice. As summarised above, through ownership alone, the new self-voice was perceived as self-relevant and prioritised perceptually. However, I extended this exploration in Experiment 3 to ask whether personally choosing the new self-voice would make it more self-relevant, and so increase the bias towards it. The results showed that perception was influenced by personal choice, such that the perceptual distance between the self-voice and friend-voice was larger in the group of participants who had personally chosen the self-voice. However, this result may not be wholly underpinned by the act of choosing the self-voice. The degree of prioritisation for the personally chosen self-voice was equivalent to the prioritisation for the assigned self-voice. Rather, the increased perceptual distance may be driven by the friend-voice being perceived as less self-relevant – and processed similarly to the other-voice – because it had been rejected as a potential self-voice. Indeed, it was deprioritised by the participants who had rejected it as a self-voice, relative to the control group who had not been given the opportunity to do so. In this instance, there was not a significant difference between bias for the friend-voice or the other-voice.

Previous literature has shown that bias – at least in memory – is maximised for items that are both self-chosen and self-owned. However, the bias is diminished for items that are self-chosen but given away to another to own. Thus, it is not the act of choosing per se that makes a stimulus more self-relevant. Rather, the role of choice in biasing our cognition may be that it affords the opportunity to reject items for the self and, in so doing, become more negatively biased towards the items that are not self-owned. Here then, the relative deprioritisation of rejected voices may be driven by other-derogation. Experiment 10 was the only other experiment within this thesis in which participants were given a choice of self-voice and a full perceptual matching paradigm was used to assess prioritisation. In Experiment 10, the voices that participants rejected as a self-voice were not then assigned to any other identity. Thus, the relative differences that I report between the self-voice and other voice identities could only be driven by bias for self and were not confounded by derogation for the other identities. In removing rejected voices from the experiment, the results from Experiment 10 show that participants attributed a small bias to the friend-voice relative to the other-voice, which is typical for the perceptual matching paradigm. This result supports the hypothesis that it is only when stimuli that have been rejected as self-stimuli are included as belonging to the other identities, that they become subject to other-derogation. Further studies will be needed to test the validity of this hypothesis and the effect of other-derogation on perceptual bias.

It is possible then, that the role of choice in integrating a new voice into the self-concept and, perhaps, the importance of choice at all, is that it enables participants to decide what is and is not self. Self-bias is attributed to stimuli that are deemed 'self' and this bias is not any more enhanced by having chosen it to be 'self', compared to having been given it. However, in choosing what is not self – by rejecting certain voices – those voices become relatively less self-relevant and were deprioritised as a result.

7.3 The Role of Agency

This thesis also examined the role of agency in a voice becoming incorporated into the self-concept. Our own voice is a signal we can use intentionally and flexibly to represent the self and communicate with others. Indeed, we have extensive control and agency over how we use our voice; to whisper, to laugh, or to shout, for instance. It was critical then to assess whether agency over the use of a new self-voice may cause that signal to be perceived as more self-relevant. Specifically, I asked whether experience of using a new, synthesised, self-owned voice modulated the extent to which it was perceptually prioritised relative to other voices. Second, I investigated whether having agency in using the voice would be reflected in the degree to which participants experienced a *sense of agency* over it.

However, neither perceptual prioritisation nor measures of sense of agency were influenced by participants' experience of using their new voice in a social context. The self-voice was prioritised relative to the other voice identities regardless of whether or not participants had had agency in using the voice. Similarly, participants' interval estimates – used as an implicit measure of agency – were significantly shorter for intervals terminating in the self-voice relative to the other-voice, but unaffected by prior experience of using the self-voice. Taken together, this shows that participants were sensitive to – and influenced by – what was self-relevant and what was not. However, communicative use of a self-voice did not, here at least, make a voice be deemed any more self-relevant.

The lack of effect of self-producing the voice suggests that the role of agency in what becomes 'self' is less important than the role of ownership. Indeed, it was through participant's ownership of one voice over another that the voices became differentially self-relevant. In turn, it was this difference in self-relevance that consistently affected the measures within the temporal binding and perceptual matching tasks. It is clear then, that having agency in producing a voice did not increase the degree to which participants felt they owned the voice; perceptual prioritisation did not increase even after having agency in using the voice. Conversely, owning the voice may affect participants' sense of agency. Specifically, participants' perception of time was compressed when they heard the self-voice outcome relative to an other-voice outcome and this effect was replicated across all experiments using the temporal binding task within this thesis. Still, future work must elucidate whether ownership affects sense of agency *per se*, or rather participants' temporal judgements more broadly, which are widely used as a measure of agency.

7.4 Implications

This thesis has significantly contributed to our understanding of what it is that can make an external voice, 'mine'; fundamentally, it is the fact that we possess a voice that causes it to be perceived as 'self'.

Indeed, across the studies presented within this thesis, it has become apparent that an unfamiliar voice may become self-relevant through ownership. This is in spite of ownership being transient and arbitrary. The voice can become self-relevant in this way very quickly, after only brief exposure to the stimulus. Thereafter, that stimulus is perceived as self-relevant such that it is prioritised in perception and people accept greater causality over it. It is particularly evident within this thesis that this bias for a new self-voice is applied irrespective of the voice's precise properties, as long as it is self-owned and, thus, self-relevant. The bias for whichever new voice was self-owned was replicated across all experiments, regardless of any of the experimental factors that were additionally manipulated. Thus, it may be that once a stimulus is explicitly tagged as self-relevant, it biases perception. This contributes to our understanding of the flexibility of the self-concept and the extent to which it can be extended to include an other-related signal. Moreover, what is 'self' can be quickly modulated via ownership and, thereafter, the increased self-relevance of the voice widely influences processing of that voice.

It appears that the preferential processing afforded to a voice is not a function of *how* self-relevant it is, only that it *is* self-relevant. After a voice had become self-relevant through ownership, it was not possible to make that new self-voice any *more* self-relevant, either through choice or agency or by the properties of the voice itself, such as how representative of the self it is. Thus, bias may be attributed in a binary way, such that self-relevance acts as a way to tag what categorically needs prioritising, relative to what is not self-relevant and should not be prioritised. This may be functionally beneficial, for example in perceptual processing in which we need to quickly differentiate what is – and is not – self-relevant. Thus, amongst a busy sensory environment, the attribution of bias as a function of self-relevance provides a means through which incoming stimuli can be distinguished as either needing a level of priority, or not.

This body of work also has positive implications for AAC users. The results suggest that makes a voice 'mine' is the basic possession of a voice. Thus, a new self-voice, of which a person attains ownership, should be prioritised perceptually as self-relevant and become a voice that they experience a greater sense of agency over relative to other voices. Moreover, the current results also support the case for choosing one's new auditory identity or, at least, having the opportunity to reject voices as the new self-voice, which may lead to greater prioritisation of the chosen self-voice.

Although having agency over producing the self-voice through the collaborative drawing game did not modulate how it was processed in this thesis, this may be different in populations of people who currently use, or will need to use, a new voice via a technological interface. Within my studies, agency and control over the voice was afforded by typing text to be said aloud in a new voice, using text-to-speech technology. This means of speech production more closely mirrors the speech production that is enabled using Augmentative and Alternative Communication (AAC) devices. However, none of the participants recruited in these studies had vocal impairments and, as such, they would be more accustomed to experiencing agency over their self-voice by using their own vocal apparatus. It remains possible that the self-relevance of a new voice may be increased if it signifies a means through which participants can regain agency in self-producing speech. Future work could recruit different populations of people (i.e., those with MND and users of AAC devices) to assess how the opportunity to have agency over a new, synthetically-generated voice may differentially influence how that voice is experienced and perceived.

7.5 Limitations

Considering the participant population tested within the experiments of this thesis, there are several limitations to note. The first is that these studies examine the perception of a self-voice that was only heard in playback and never during live speech. Participants who can still use their own voice to communicate, such as those tested in the present research, will have prior experience of hearing their own voice in playback (e.g., a recorded interview, a voice memo), but this is not a frequent experience, and it can be qualitatively different to hear one's voice only "in the air" rather than also "in the head" (via bone and tissue conduction). Future studies may look to incorporate the new self-voice into a speech production paradigm to test this.

Similarly, studies thus far have not been able to examine sense of agency over the biological production of a new self-voice, presumably due to technological constraints.

However, in a previous study by Zheng, MacDonald, Munhall, & Johnsrude (2017), a stranger's voice was presented as auditory feedback to a participant's own speech production. When the stranger's voice was temporally and phonetically congruent with the participants' speech production, it was perceived as the self-voice. Similarly, Franken, Hartsuiker, Johansson, Hall, & Lind, (2021) conducted a study in which they pitch-shifted participants' voices in near-real time, using an altered feedback paradigm. The authors measured the participants' sense of agency according to how far they adapted their own voice production to hearing the altered feedback; hypothesising that only outcomes that participants had accepted as having been self-generated would drive an adaptive response. Within both studies, the authors concluded that recognition of what is 'self' is surprisingly flexibly and context dependent, and that another's voice can quickly become perceived as self through such manipulations. However, within these studies, the acceptance of a different voice as 'self' relies on illusion rather than explicit integration of a new voice. Therefore, further work is needed to assess the extent to which people can accept externally-generated sounds as being 'self' according to how these are produced (e.g. via altered feedback versus in playback).

A second limitation of this research is the fact that the study of self-bias is subject to differences in each 'self'. There are large individual differences in subjects across the studies. Mezulis, Abramson, Hyde, and Hankin (2004) have previously run a meta-analysis of individual, developmental, and cultural differences in the attribution of self-bias. Their review concluded that self-bias is robust and pervasive in the general population but is subject to variability according to age, culture, and psychopathy. The eligibility criteria I used for these studies typically demanded participants to be in the age group of 18-40, with UK nationality and English as their first language, and with no visual or hearing impairments. However, accounting for all the possible factors that may contribute to an individual's attribution of self-bias was beyond the remit of this thesis and would require a fuller investigation of individual differences. Thus, this thesis does not focus on how the 'self' influences the attribution of bias, but rather how bias is attributed according to the characteristics of the to-be-associated voice. It will be important for future work to examine these in tandem and explore how one's perception is shaped and guided not only by how the self can be extended, but also by what is already ‘self’.

Relatedly, the studies reported within this thesis have not accounted for the extent to which individual performance is affected by differences in voice perception ability. Aglieri, Watson, Pernet, Latinus, Garrido & Belin, (2016; see also Lavan, Burston, & Garrido, 2019) have previously shown that there is substantial variability in listeners’ ability to discriminate unfamiliar voices. Given that the prioritisation of a new self-voice relies on the fact that it has been identified as the self-voice, it may be that some of the variation in self-bias is underpinned by variation in voice perception ability.

Finally, this thesis has been unable to fully disentangle how the effect of ownership and of self-association interacts with the temporal binding paradigm. In each interval estimation task, participants’ estimates were influenced by whether the voice outcome was self-associated or not. However, this result was also present in the passive condition which, importantly, demanded no agentic action from the participant. Thus, it is unclear whether self-association does affect agency, as suggested by the presence of an effect in active trials, or whether it affects temporal perception more broadly. Further studies, briefly outlined below, may help to resolve these limitations, corroborate the current results, and further our understanding.

7.6 Future studies

This body of work presents an extensive examination of how a new voice may become relevant to the self-concept according to behavioural, mainly implicit, measures. Further studies may help to corroborate these findings through different psychophysics techniques and neuroimaging work.

In the first instance, it will be important to clarify other factors to which the temporal binding paradigm is sensitive. This thesis has demonstrated that participants’ temporal judgements of interval durations are influenced by self-association but that this does not necessarily interact with participants’ sense of agency. Further studies must elucidate a more precise understanding of what factors drive interval estimates, whether these factors are exclusively modulated under conditions of agency and, thus, whether interval estimates can be conclusively interpreted as reflecting only a sense of agency.

In order to reconcile the consistent presence of an effect of self-association (i.e., whether a voice was self-owned or other-owned) in passive trials as well as active trials, I hypothesised that the quicker perceptual processing of the self-owned voice may be what impacts perception of the interval duration. Specifically, that after a starting cue, the time between that cue and a voice is postdictively shortened according to how quickly the voice is recognised. This hypothesis was motivated by the fact that the duration of the cue-tone could similarly influence how long the subsequent interval was perceived to last; shorter cue-tones led to the interval being perceived as correspondingly shorter. However, this hypothesis is untested and not fully supported by the current results. In Experiment 6, in which participants were presented with their true self-voice as 'self', their reaction times to this voice in the perceptual matching paradigm were significantly quicker than reaction times to the friend and stranger voice identities; they also responded faster to this true self-voice than control participants who were assigned it as a completely novel self-voice in the experiment. Despite this quicker recognition of the true self-voice, however, the interval preceding it in the interval estimation task was not perceived as any shorter than intervals preceding a new self-voice.

Moving forward, it may be possible to disentangle the influence of self-association on sense of agency by using a different measure of agency, such as the original Libet clock method (Libet, Gleason, Wright, and Pearl, 1983). This method requires participants to use a visible clock-face with a rotating hand to estimate the time at which an action is made and, separately, the time at which an outcome occurs. Then, rather than a passive condition, participants estimate the times at which these events occur in isolation (i.e., the action when it is not followed by an outcome, and the outcome without its being self-generated). With this method, it may be possible to determine whether participants experience a greater sense of agency over their *actions* that result in a self-voice outcome in comparison to their actions that result in an other-voice outcome. Moreover, by measuring estimates of when the outcome occurred in isolation, it may be possible to see whether the effect of self-association is present in the absence of action. This method, in conjunction with the current results, may provide clarity on how self-relevance interacts with participants' sense of agency over their actions specifically, as separated from their temporal perception of the outcome's onset.

Future work may also serve to elucidate how far a new self-voice becoming self-relevant is analogous to it becoming 'self'. It is clear within these studies that people can quickly process a new voice as *self-relevant,* as shown by robust behavioural results. However, further work could be done to determine the extent to which the new voice is processed as self. Payne et al. (2017) demonstrated that participants could process the face of another as self-relevant when they owned it. However, they also found that participants' physical representation of the self-face was not changed. It will be important for future work to assess how far a new self-voice has been integrated within the self-concept and this could, in the first instance, be explored with a similar task as Payne et al. (2017). Specifically, by morphing the true self-voice with a new self-owned voice and assessing the point at which the merged voice is deemed more self than other, or more other than self. A change in representation would be indicated by the point of self-other equivalence changing after the new voice has become self-relevant.

Neuroimaging studies could also help to corroborate the extent to which a new voice has been integrated into the concept of self, and further, into the representation of self. Previous functional MRI studies have indicated that the ventral medial prefrontal cortex is more actively engaged for self-relevant stimuli and this brain region has been linked to self-representation. Thus, an fMRI study that investigates the neural correlates of the self-prioritisation effect in voices could further the current work. Specifically, an fMRI study could investigate differential brain activity when the new self-owned voice is presented compared to when the voices belonging to the friend and other identities are presented. Such a study would require looking at patterns of activation in the ventral medial pre-frontal cortex (vmPFC), the left posterior superior temporal sulcus (lpSTS) and functional coupling between the two. If the new self-owned voice really has been processed as 'self' one might expect to see greater functional connectivity changes between the vmPFC and lpSTS for the self-owned voice relative to the other voices. These results could be further validated by testing for correlations between brain responses and the observed bias for the self-voice in the behavioural matching task (measuring RTs and proportion accuracy).

7.7 Conclusion

Overall then, this body of work aimed to investigate the possibility of incorporating a new voice into the self-concept and, by way of this, further our understanding of what it is that makes a voice ‘mine’. This research demonstrates, for the first time, that a new voice can be incorporated into the self-concept and become self-relevant; a voice, moreover, that is the inherent biological property of another. This is achieved predominantly thorough taking ownership of it, which fundamentally changes the relation between the self and that stimulus such that it is made self-relevant and afforded bias. This shows the flexibility of the self-concept, both in terms of the information that comprises it and in how quickly that information is adjusted. Thus, in beginning to answer what it is that makes a voice ‘mine’: simply having a voice is critical to perceiving that voice as self-relevant and to experiencing a sense of agency over it.

References

Allan, K., Morson, S., Dixon, S., Martin, D., & Cunningham, S. (2017). Simulation-Based Mentalizing Generates a “proxy” Self-Reference Effect in Memory. *Quarterly Journal Of Experimental Psychology*, *70*(6), 1074-1084. doi: 10.1080/17470218.2016.1209532

Allard, E., & Williams, D. (2008). Listeners’ perceptions of speech and language disorders. *Journal Of Communication Disorders*, *41*(2), 108-123. doi: 10.1016/j.jcomdis.2007.05.002

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends In Cognitive Sciences*, *4*(7), 267-278. doi: 10.1016/s1364-6613(00)01501-1

Ally, B., & Budson, A. (2007). The worth of pictures: Using high density event-related potentials to understand the memorial power of pictures and the dynamics of recognition memory. *Neuroimage*, *35*(1), 378-395. doi: 10.1016/j.neuroimage.2006.11.023

Alós-Ferrer, C., & Shi, F. (2012). Choice-Induced Preference Change: In Defense of the Free-Choice Paradigm. *SSRN Electronic Journal*. doi: 10.2139/ssrn.2062507

Ariely, D., & Norton, M. (2008). How actions create – not just reveal preferences. *Trends In Cognitive Sciences*, 12(1), 13-16. doi: 10.1016/j.tics.2007.10.008

Audacity Team (2019). Audacity(R): Free Audio Editor and Recorder [Computer application]. Version 2.3.2 Retrieved March 2019 from https://audacityteam.org/.

Banda, D., & Alzrayer, N. (2018). This meta-analysis provides some evidence to support the use of high-tech AAC interventions to improve social-communication skills in individuals with intellectual and developmental disabilities. *Evidence-Based Communication Assessment And Intervention*, *12*(3), 73-76. doi: 10.1080/17489539.2018.1482633

Banai, B., Laustsen, L., Banai, I. and Bovan, K., 2018. Presidential, But Not Prime Minister, Candidates With Lower Pitched Voices Stand a Better Chance of Winning the Election in Conservative Countries. *Evolutionary Psychology*, 16(2), p.147470491875873.

Barlas, Z., & Kopp, S. (2018). Action Choice and Outcome Congruency Independently Affect Intentional Binding and Feeling of Control Judgments. *Frontiers In Human Neuroscience*, *12*. doi: 10.3389/fnhum.2018.00137

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. *Journal of Statistical Software*, **67**, 1– 23. https://doi.org/10.18637/jss.v067.i01

Beardsworth, T., & Buckner, T. (1981). The ability to recognize oneself from a video recording of one's movements without seeing one's body. *Bulletin Of The Psychonomic Society*, *18*(1), 19-22. doi: 10.3758/bf03333558

Belin, P., Bestelmeyer, P., Latinus, M., & Watson, R. (2011). Understanding Voice Perception. *British Journal Of Psychology*, *102*(4), 711-725. doi: 10.1111/j.2044-8295.2011.02041.x

Belin, P., Fecteau, S., & Bédard, C. (2004). Thinking the voice: neural correlates of voice perception. *Trends In Cognitive Sciences*, *8*(3), 129-135. doi: 10.1016/j.tics.2004.01.008

Belk, R. (1988). Possessions and the Extended Self. *Journal Of Consumer Research*, *15*(2), 139. doi: 10.1086/209154

Bendapudi, N., & Leone, R. (2003). Psychological Implications of Customer Participation in Co-Production. *Journal Of Marketing*, *67*(1), 14-28. doi: 10.1509/jmkg.67.1.14.18592

Benoit, R.G., Gilbert, S.J., Volle, E., & Burgess, P.W. (2010). When I think about me and simulate you: medial rostral prefrontal cortex and self-referential processes. *NeuroImage*, **50**, 1340– 1349. https://doi.org/10.1016/j.neuroimage.2009.12.091

Bentley, S., Peters, K., Haslam, S., & Greenaway, K. (2019). Construction at Work: Multiple Identities Scaffold Professional Identity Development in Academia. *Frontiers In Psychology*, *10*. doi: 10.3389/fpsyg.2019.00628

Bestelmeyer, P., Belin, P., & Ladd, D. (2014). A Neural Marker for Social Bias Toward In-group Accents. *Cerebral Cortex*, *25*(10), 3953-3961. doi: 10.1093/cercor/bhu282

Beukelman, D., Fager, S., & Nordness, A. (2011). Communication Support for People with ALS. *Neurology Research International*, *2011*, 1-6. doi: 10.1155/2011/714693

Beyer, F., Sidarus, N., Bonicalzi, S., & Haggard, P. (2016). Beyond self-serving bias: diffusion of responsibility reduces sense of agency and outcome monitoring. *Social Cognitive And Affective Neuroscience*, *12*(1), 138-145. doi: 10.1093/scan/nsw160

Bigelow, J., & Poremba, A. (2014). Achilles' Ear? Inferior Human Short-Term and Recognition Memory in the Auditory Modality. *Plos ONE*, *9*(2), e89914. doi: 10.1371/journal.pone.0089914

Biggio, M., Bisio, A., Avanzino, L., Ruggeri, P., & Bove, M. (2020). Familiarity with a Tool Influences Peripersonal Space and Primary Motor Cortex Excitability of Muscles Involved in Haptic Contact. *Cerebral Cortex Communications*, *1*(1). doi: 10.1093/texcom/tgaa065

v. Békésy, G. (1954). Note on the Definition of the Term: Hearing by Bone Conduction. *The Journal Of The Acoustical Society Of America*, *26*(1), 106-107. doi: 10.1121/1.1907278

Boersma, P., & Weenink, D. (2020). Praat: doing phonetics by computer [Computer program]. Version 6.1.26, retrieved 5 October 2019 from http://www.praat.org

Bortolon, C., & Raffard, S. (2018). Self-face advantage over familiar and unfamiliar faces: A three-level meta-analytic approach. *Psychonomic Bulletin & Review*, *25*(4), 1287-1300. doi: 10.3758/s13423-018-1487-9

Botvinick, M., & Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature*, *391*(6669), 756-756. doi: 10.1038/35784

Braun, N., Debener, S., Spychala, N., Bongartz, E., Sörös, P., Müller, H., & Philipsen, A. (2018). The Senses of Agency and Ownership: A Review. *Frontiers In Psychology*, *9*. doi: 10.3389/fpsyg.2018.00535

Brédart, S. (2016). A Self-Reference Effect on Memory for People: We Are Particularly Good at Retrieving People Named Like Us. *Frontiers In Psychology*, *7*. doi: 10.3389/fpsyg.2016.01751

Brozzoli, C., Pavani, F., Urquizar, C., Cardinali, L., & Farnè, A. (2009). Grasping actions remap peripersonal space. *Neuroreport*, *20*, 913–917. https://doi.org/10.1097/WNR.0b013e32832c0b9b

Brozzoli, C., Makin, T. R., Cardinali, L., Holmes, N. P., and Farnè, A. (eds) (2011). "Peripersonal Space: a multisensory interface for body-object interactions," in *The Neural Bases of Multisensory Processes* (London: Taylor & Francis), 449–466.

Buehner, M. J. (2012). Understanding the past, predicting the future causation, not intentional action, is the root of temporal binding. *Psychological science*, *23*, 1490- 1497.

Candini, M., Avanzi, S., Cantagallo, A., Zangoli, M., Benassi, M., Querzani, P., Lotti, E., Iachini, T. and Frassinetti, F., 2018. The lost ability to distinguish between self and other voice following a brain lesion. *NeuroImage: Clinical*, 18, pp.903-911.

Caspar, E., Christensen, J., Cleeremans, A., & Haggard, P. (2016). Coercion Changes the Sense of Agency in the Human Brain. *Current Biology*, *26*(5), 585-592. doi: 10.1016/j.cub.2015.12.067

Cave, R., & Bloch, S. (2020). Voice banking for people living with motor neurone disease: Views and expectations. *International Journal Of Language & Communication Disorders*, *56*(1), 116-129. doi: 10.1111/1460-6984.12588

Chambon, V., & Haggard, P. (2012). Sense of control depends on fluency of action selection, not motor performance. *Cognition*, *125*(3), 441-451. doi: 10.1016/j.cognition.2012.07.011

Cheatwell Games. (2004). Up A Bit [Board game]. Hampshire, UK: Outset Media

Chiu, C., Ho, H., & Tollenaar, M. (2021). Relational self-evaluations in dissociation: Implicit self-rejection?. *Psychological Trauma: Theory, Research, Practice, And Policy*. doi: 10.1037/tra0001017

Ciardo, F., Beyer, F., De Tommaso, D., & Wykowska, A. (2020). Attribution of intentional agency towards robots reduces one's own sense of agency. *Cognition*, *194*, 104109. doi: 10.1016/j.cognition.2019.104109

Clark, A. (2007). Soft selves and ecological control. In D. Ross, D. Spurrett, H. Kincaid & G. L. Stephens (eds.), *Distributed Cognition and the Will: Individual Volition and Social Context*. MIT Press. pp. 101--122.

Cloutier, J., & Neil Macrae, C. (2008). The feeling of choosing: Self-involvement and the cognitive status of things past. *Consciousness And Cognition*, *17*(1), 125-135. doi: 10.1016/j.concog.2007.05.010

Cohen, M., Horowitz, T., & Wolfe, J. (2009). Auditory recognition memory is inferior to visual recognition memory. *Proceedings Of The National Academy Of Sciences*, *106*(14), 6008-6010. doi: 10.1073/pnas.0811884106

Connor, N., Cohen, S., Theis, S., Thibeault, S., Heatley, D., & Bless, D. (2008). Attitudes of Children With Dysphonia. *Journal Of Voice*, *22*(2), 197-209. doi: 10.1016/j.jvoice.2006.09.005

Constable, M., Welsh, T., Huffman, G., & Pratt, J. (2018). I before U: Temporal order judgements reveal bias for self-owned objects. *Quarterly Journal Of Experimental Psychology*, *72*(3), 589-598. doi: 10.1177/1747021818762010

Conway, M., & Pleydell-Pearce, C. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, *107*(2), 261-288. doi: 10.1037/0033-295x.107.2.261

Cravo, A., Claessens, P., & Baldo, M. (2009). Voluntary action and causality in temporal binding. *Experimental Brain Research*, *199*(1), 95-99. doi: 10.1007/s00221-009-1969-0

Cunningham, S., Brady-Van den Bos, M., & Turk, D. (2011). Exploring the effects of ownership and choice on self-memory biases. *Memory*, *19*(5), 449-461. doi: 10.1080/09658211.2011.584388

Cunningham, S., Brady-Van den Bos, M., & Turk, D. (2011). Exploring the effects of ownership and choice on self-memory biases. *Memory*, *19*(5), 449-461. doi: 10.1080/09658211.2011.584388

Cunningham, S., Turk, D., Macdonald, L., & Neil Macrae, C. (2008). Yours or mine? Ownership and memory. *Consciousness And Cognition*, *17*(1), 312-318. doi: 10.1016/j.concog.2007.04.003

de Haan, A. M., Smit, M., Van der Stigchel, S., and Dijkerman, H. C. (2016). Approaching threat modulates visuotactile interactions in peripersonal space. *Exp. Brain Res.* 234, 1875–1884. doi: 10.1007/s00221-016-4571-2

Denny, B., Kober, H., Wager, T., & Ochsner, K. (2012). A Meta-analysis of Functional Neuroimaging Studies of Self- and Other Judgments Reveals a Spatial Gradient for Mentalizing in Medial Prefrontal Cortex. *Journal Of Cognitive Neuroscience*, *24*(8), 1742-1752. doi: 10.1162/jocn_a_00233

Desantis, A., Hughes, G., & Waszak, F. (2012). Intentional Binding Is Driven by the Mere Presence of an Action and Not by Motor Prediction. *Plos ONE*, *7*(1), e29557. doi: 10.1371/journal.pone.0029557

Desantis, A., Roussel, C., & Waszak, F. (2011). On the influence of causal beliefs on the feeling of agency. *Consciousness And Cognition*, *20*(4), 1211-1220. doi: 10.1016/j.concog.2011.02.012

Devue, C., & Brédart, S. (2011). The neural correlates of visual self-recognition. *Consciousness And Cognition*, *20*(1), 40-51. doi: 10.1016/j.concog.2010.09.007

Dhakal V, Feit AM, Kristensson PO, Oulasvirta A (2018) Observations on Typing from 136 Million Keystrokes, Association for Computing Machinery, New York, NY, USA, pp 1–12. https://doi.org/10.1145/3173574.3174220

Di, X., Zhu, S., Jin, H., Wang, P., Ye, Z., & Zhou, K. et al. (2012). Altered Resting Brain Function and Structure in Professional Badminton Players. *Brain Connectivity*, *2*(4), 225-233. doi: 10.1089/brain.2011.0050

DiQuattro, N., & Geng, J. (2011). Contextual Knowledge Configures Attentional Control Networks. *Journal Of Neuroscience*, *31*(49), 18026-18035. doi: 10.1523/jneurosci.4040-11.2011

Dogge, M., Schaap, M., Custers, R., Wegner, D., & Aarts, H. (2012). When moving without volition: Implied self-causation enhances binding strength between involuntary actions and effects. *Consciousness And Cognition*, *21*(1), 501-506.doi: 10.1016/j.concog.2011.10.014

Dohle, S., Rall, S., & Siegrist, M. (2014). I cooked it myself: Preparing food increases liking and consumption. *Food Quality And Preference*, *33*, 14-16. doi: 10.1016/j.foodqual.2013.11.001

Dommer, S., & Swaminathan, V. (2013). Explaining the Endowment Effect through Ownership: The Role of Identity, Gender, and Self-Threat. *Journal Of Consumer Research*, *39*(5), 1034-1050. doi: 10.1086/666737

Dux, P., & Marois, R. (2009). The attentional blink: A review of data and theory. *Attention, Perception & Psychophysics*, *71*(8), 1683-1700. doi: 10.3758/app.71.8.1683

Eagleman, D., & Holcombe, A. (2002). Causality and the perception of time. *Trends In Cognitive Sciences*, *6*(8), 323-325. doi: 10.1016/s1364-6613(02)01945-9

Ebert, J., & Wegner, D. (2010). Time warp: Authorship shapes the perceived timing of actions and events. *Consciousness And Cognition*, *19*(1), 481-489. doi: 10.1016/j.concog.2009.10.002

Engbert, K., Wohlschläger, A., & Haggard, P. (2008). Who is causing what? The sense of agency is relational and efferent-triggered. *Cognition*, *107*(2), 693-704. doi: 10.1016/j.cognition.2007.07.021

Engbert, K., Wohlschläger, A., Thomas, R., & Haggard, P. (2007). Agency, subjective time, and other minds. *Journal Of Experimental Psychology: Human Perception And Performance*, *33*(6), 1261-1268. doi: 10.1037/0096-1523.33.6.1261

Enock, F., Hewstone, M., Lockwood, P., & Sui, J. (2020). Overlap in processing advantages for minimal ingroups and the self. *Scientific Reports*, *10*(1). doi: 10.1038/s41598-020-76001-9

Enock, F., Sui, J., Hewstone, M., & Humphreys, G. (2018). Self and team prioritisation effects in perceptual matching: Evidence for a shared representation. *Acta Psychologica*, *182*, 107-118. doi: 10.1016/j.actpsy.2017.11.011

Falbén, J., Golubickis, M., Balseryte, R., Persson, L., Tsamadi, D., Caughey, S., & Neil Macrae, C. (2019). How prioritized is self-prioritization during stimulus processing?. *Visual Cognition*, *27*(1), 46-51. doi: 10.1080/13506285.2019.1583708

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G□Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175–191. doi: 10.3758/BF03193146

Forgas, J.P., & Williams, K.D. (Eds.). (2003). The Social Self: Cognitive, Interpersonal and Intergroup Perspectives (1st ed.). Psychology Press. https://doi.org/10.4324/9781315800516

Fourkas, A., Bonavolonta, V., Avenanti, A., & Aglioti, S. (2008). Kinesthetic Imagery and Tool-Specific Modulation of Corticospinal Representations in Expert Tennis Players. *Cerebral Cortex*, *18*(10), 2382-2390. doi: 10.1093/cercor/bhn005

Franke, N., Keinz, P., & Steger, C. (2009). Testing the Value of Customization: When Do Customers Really Prefer Products Tailored to Their Preferences?. *Journal Of Marketing*, *73*(5), 103-121. doi: 10.1509/jmkg.73.5.103

Franken, M., Hartsuiker, R., Johansson, P., Hall, L. and Lind, A., 2021. Speaking with an alien voice: Flexible sense of agency during vocal production. *Journal of Experimental Psychology: Human Perception and Performance*,.

Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends In Cognitive Sciences*, *4*(1), 14-21. doi: 10.1016/s1364-6613(99)01417-5

Gardiner, J., Gregg, V., Mashru, R., & Thaman, M. (2001). Impact of encoding depth on awareness of perceptual effects in recognition memory. *Memory & Cognition*, *29*(3), 433-440. doi: 10.3758/bf03196394

Gerson, S., Meyer, M., Hunnius, S., & Bekkering, H. (2017). Unravelling the contributions of motor experience and conceptual knowledge in action perception: A training study. *Scientific Reports*, *7*(1). doi: 10.1038/srep46761

Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. In H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of accommodation: Developments in applied sociolinguistics* (pp. 1–68). Cambridge University Press; Editions de la Maison des Sciences de l'Homme. https://doi.org/10.1017/CBO9780511663673.001

de Greck, M., Rotte, M., Paus, R., Moritz, D., Thiemann, R., & Proesch, U. et al. (2008). Is our self based on reward? Self-relatedness recruits neural activity in the reward system. *Neuroimage*, *39*(4), 2066-2075. doi: 10.1016/j.neuroimage.2007.11.006

Gregg, A., Mahadevan, N., & Sedikides, C. (2017). The SPOT effect: people spontaneously prefer their own theories. *Quarterly Journal of Experimental Psychology*, **70**, 996– 1010. https://doi.org/10.1080/17470218.2015.1099162

Golubickis, M., Falbén, J., Ho, N., Sui, J., Cunningham, W., & Neil Macrae, C. (2020). Parts of me: Identity-relevance moderates self-prioritization. *Consciousness And Cognition*, *77*, 102848. doi: 10.1016/j.concog.2019.102848

Grady, C., McIntosh, A., Rajah, M., & Craik, F. (1998). Neural correlates of the episodic encoding of pictures and words. *Proceedings Of The National Academy Of Sciences*, *95*(5), 2703-2708. doi: 10.1073/pnas.95.5.2703

Guldner, S., Nees, F., & McGettigan, C. (2020). Vocomotor and Social Brain Networks Work Together to Express Social Traits in Voices. *Cerebral Cortex*, *30*(11), 6004-6020. doi: 10.1093/cercor/bhaa175

Gusnard, D., Akbudak, E., Shulman, G., & Raichle, M. (2001). Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proceedings Of The National Academy Of Sciences*, *98*(7), 4259-4264. doi: 10.1073/pnas.071043098

Haering, C., & Kiesel, A. (2011). Time in action contexts: learning when an action effect occurs. *Psychological Research*, *76*(3), 336-344. doi: 10.1007/s00426-011-0341-8

Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience*, *18*(4), 196-207. doi: 10.1038/nrn.2017.14

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, *5*(4), 382-385. doi: 10.1038/nn827

Hamlin, J., Mahajan, N., Liberman, Z., & Wynn, K. (2013). Not Like Me = Bad. *Psychological Science*, *24*(4), 589-594. doi: 10.1177/0956797612457785

Hauf, P., Elsner, B., & Aschersleben, G. (2004). The role of action effects in infants? action control. *Psychological Research*, *68*(2-3), 115-125. doi: 10.1007/s00426-003-0149-2

Heersmink, R. (2020). Varieties of the extended self. *Consciousness And Cognition*, *85*, 103001. doi: 10.1016/j.concog.2020.103001

Holmes, E., To, G., & Johnsrude, I. (2021). How Long Does It Take for a Voice to Become Familiar? Speech Intelligibility and Voice Recognition Are Differentially Sensitive to Voice Training. *Psychological Science*, *32*(6), 903-915. doi: 10.1177/0956797621991137

Hu, C., Lan, Y., Macrae, C., & Sui, J. (2020). Good Me Bad Me: Prioritization of the Good-Self During Perceptual Decision-Making. *Collabra: Psychology*, *6*(1). doi: 10.1525/collabra.301

Hu, C., Wang, Q., Short, L., & Fu, G. (2012). Speech Spectrum's Correlation with Speakers' Eysenck Personality Traits. *Plos ONE*, *7*(3), e33906. doi: 10.1371/journal.pone.0033906

Huang, Y., Wang, L., & Shi, J. (2009). When do objects become more attractive? The individual and interactive effects of choice and ownership on object evaluation. *Personality and Social Psychology Bulletin*, **35**, 713– 722. https://doi.org/10.1177/0146167209333046

Huang, Y., & Wu, Y. (2016). Ownership Effect Can Be a Result of Other-Derogation: Evidence from Behavioral and Electrophysiological Studies. *PLOS ONE*, *11*(11), e0166054. doi: 10.1371/journal.pone.0166054

Huffman, G., & Brockmole, J. (2020). Attentional selection is biased towards controllable stimuli. *Attention, Perception, & Psychophysics*, *82*(5), 2558-2569. doi: 10.3758/s13414-020-02004-3

Hughes, S., & Harrison, M. (2013). I like my voice better: self-enhancement bias in perceptions of voice attractiveness. *Perception*, **42**, 941– 949. https://doi.org/10.1068/p7526

Hughes, S., & Nicholson, S. (2010). The processing of auditory and visual recognition of self-stimuli. *Consciousness And Cognition*, *19*(4), 1124-1134. doi: 10.1016/j.concog.2010.03.001

Humphreys, G., & Sui, J. (2016). Attentional control and the self: The Self-Attention Network (SAN). *Cognitive Neuroscience*, *7*(1-4), 5-17. doi: 10.1080/17588928.2015.1044427

Imaizumi, S., & Tanno, Y. (2019). Intentional binding coincides with explicit sense of agency. *Consciousness And Cognition*, *67*, 1-15. doi: 10.1016/j.concog.2018.11.005

Jiang, X., & Pell, M. (2015). On how the brain decodes vocal cues about speaker confidence. *Cortex*, *66*, 9-34. doi: 10.1016/j.cortex.2015.02.002

Jiang, X., & Pell, M. (2017). The sound of confidence and doubt. *Speech Communication*, *88*, 106-126. doi: 10.1016/j.specom.2017.01.011

Jiang, X., Gossack-Keenan, K., & Pell, M. (2019). To believe or not to believe? How voice and accent information in speech alter listener impressions of trust. *Quarterly Journal Of Experimental Psychology*, *73*(1), 55-79. doi: 10.1177/1747021819865833

Johns, M., Arviso, L., & Ramadan, F. (2011). Challenges and Opportunities in the Management of the Aging Voice. *Otolaryngology–Head And Neck Surgery*, *145*(1), 1-6. doi: 10.1177/0194599811404640

Johnsrude, I., Mackey, A., Hakyemez, H., Alexander, E., Trang, H., & Carlyon, R. (2013). Swinging at a Cocktail Party. *Psychological Science*, *24*(10), 1995-2004. doi: 10.1177/0956797613482467

Jones, J., Pelham, B., Carvallo, M., & Mirenberg, M. (2004). How do I love thee? Let me count the Js: Implicit egotism and interpersonal attraction. *Journal of Personality and Social Psychology*, **87**, 665– 683. https://doi.org/10.1037/0022-3514.87.5.665

Kahneman, D., Knetsch, J., & Thaler, R. (1991). Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias. *Journal Of Economic Perspectives*, *5*(1), 193-206. doi: 10.1257/jep.5.1.193

Kaiser, U., Schreier, M., & Janiszewski, C. (2017). The Self-Expressive Customization of a Product Can Improve Performance. *Journal Of Marketing Research*, *54*(5), 816-831. doi: 10.1509/jmr.14.0293

Kanber, E., Lavan, N., & McGettigan, C. (2021). Highly accurate and robust identity perception from personally familiar voices. *Journal Of Experimental Psychology: General*. doi: 10.1037/xge0001112

Kelley, W., Macrae, C., Wyland, C., Caglar, S., Inati, S., & Heatherton, T. (2002). Finding the Self? An Event-Related fMRI Study. *Journal Of Cognitive Neuroscience*, *14*(5), 785-794. doi: 10.1162/08989290260138672

Kesebir, S., & Oishi, S. (2010). A Spontaneous Self-Reference Effect in Memory. *Psychological Science*, *21*(10), 1525-1531. doi: 10.1177/0956797610383436

Keyes, H., & Dlugokencka, A. (2014). Do I Have My Attention? Speed of Processing Advantages for the Self-Face Are Not Driven by Automatic Attention Capture. *Plos ONE*, *9*(10), e110792. doi: 10.1371/journal.pone.0110792

Kim, K., Jeon, Y., Banquer, A., & Rothschild, D. (2018). Conscious awareness of self-relevant information is necessary for an incidental self-memory advantage. *Consciousness And Cognition*, *65*, 228-239. doi: 10.1016/j.concog.2018.09.004

Kimura, M., & Yotsumoto, Y. (2018). Auditory traits of "own voice". *PLOS ONE*, *13*(6) e0199443. doi: 10.1371/journal.pone.0199443

Kirsch, W., Kunde, W., & Herbort, O. (2019). Intentional binding is unrelated to action intention. *Journal Of Experimental Psychology: Human Perception And Performance*, *45*(3), 378-385. doi: 10.1037/xhp0000612

Knoblich, G., Seigerschmidt, E., Flach, R., & Prinz, W. (2002). Authorship effects in the prediction of handwriting strokes: Evidence for action simulation during action perception. *The Quarterly Journal Of Experimental Psychology Section A*, *55*(3), 1027-1046. doi: 10.1080/02724980143000631

Hannah, L., David, C., & James, M. (2015). The Sense of Agency during Verbal Action. *Frontiers In Human Neuroscience*, *9*. doi: 10.3389/conf.fnhum.2015.217.00294

Knoblich, G., & Prinz, W. (2001). Recognition of self-generated actions from kinematic displays of drawing. *Journal Of Experimental Psychology: Human Perception And Performance*, *27*(2), 456-465. doi: 10.1037/0096-1523.27.2.456

Kreiman, J., & Sidtis, D. (2011). Voices and Listeners: Toward a Model of Voice Perception. *Acoustics Today*, *7*(4), 7. doi: 10.1121/1.3684228

Langarika-Rocafort, A., Mondragon, N., & Etxebarrieta, G. (2021). A Systematic Review of Research on Augmentative and Alternative Communication Interventions for Children Aged 6–10 in the Last Decade. *Language, Speech, And Hearing Services In Schools*, *52*(3), 899-916. doi: 10.1044/2021_lshss-20-00005

Lass, N., Ruscello, D., Stout, L., & Hoffmann, F. (1991). Peer Perceptions of Normal and Voice-Disordered Children. *Folia Phoniatrica Et Logopaedica*, *43*(1), 29-35. doi: 10.1159/000266098

Lavan, N., Burton, A., Scott, S., & McGettigan, C. (2018). Flexible voices: Identity perception from variable vocal signals. *Psychonomic Bulletin & Review*, *26*(1), 90-102. doi: 10.3758/s13423-018-1497-7

Lavan, N., Kreitewolf, J., Obleser, J., & McGettigan, C. (2021). Familiarity and task context shape the use of acoustic information in voice identity perception. *Cognition*, *215*, 104780. doi: 10.1016/j.cognition.2021.104780

Lawrence, R., & Chai, X. (2021). Self-referential encoding of source information in recollection memory. *PLOS ONE*, *16*(4), e0248044. doi: 10.1371/journal.pone.0248044

Legaspi, R., & Toyoizumi, T. (2019). A Bayesian psychophysics model of sense of agency. *Nature Communications*, *10*(1). doi: 10.1038/s41467-019-12170-0

Lenth, R. (2016). Least-squares means: the R package lsmeans. *Journal of Statistical Software*, **69**(1), 1– 33. https://doi.org/10.18637/jss.v069.i01

Leotti, L., & Delgado, M. (2011). The Inherent Reward of Choice. *Psychological Science*, *22*(10), 1310-1318. doi: 10.1177/0956797611417005

Li, S., Zhu, X., Ding, R., Ren, J., & Luo, W. (2019). The effect of emotional and self-referential contexts on ERP responses towards surprised faces. *Biological Psychology*, *146*, 107728. doi: 10.1016/j.biopsycho.2019.107728

Libet, B., Gleason, C., Wright, E., & Pearl, D. (1983). Time of conscious intention to act in relation to onset of cerebral activity (Readiness-Potential)). *Brain*, *106*(3), 623-642. doi: 10.1093/brain/106.3.623

Ma, Y., & Han, S. (2010). Why we respond faster to the self than to others? An implicit positive association theory of self-advantage during implicit face recognition. *Journal Of Experimental Psychology: Human Perception And Performance*, *36*(3), 619-633. doi: 10.1037/a0015797

MacLeod, C. (2011). I said, you said: The production effect gets personal. *Psychonomic Bulletin & Review*, *18*(6), 1197-1202. doi: 10.3758/s13423-011-0168-8

MacLeod, C., Gopie, N., Hourihan, K., Neary, K., & Ozubko, J. (2010). The production effect: Delineation of a phenomenon. *Journal Of Experimental Psychology: Learning, Memory, And Cognition*, *36*(3), 671-685. doi: 10.1037/a0018785

Mahajan, N., & Wynn, K. (2012). Origins of “Us” versus “Them”: Prelinguistic infants prefer similar others. *Cognition*, *124*(2), 227-233. doi: 10.1016/j.cognition.2012.05.003

Makwana, M., & Srinivasan, N. (2019). Self-associated stimuli produce stronger intentional binding. *Journal Of Experimental Psychology: Human Perception And Performance*, *45*(11), 1436-1442. doi: 10.1037/xhp0000687

Mama, Y., & Icht, M. (2014). Auditioning the distinctiveness account: Expanding the production effect to the auditory modality reveals the superiority of writing over vocalising. *Memory*, *24*(1), 98-113. doi: 10.1080/09658211.2014.986135

Maslowski, M., Meyer, A., & Bosker, H. (2018). Listening to yourself is special: Evidence from global speech rate tracking. *PLOS ONE*, *13*(9), e0203571. doi: 10.1371/journal.pone.0203571

Mathias, S. & von Kreigstein, K. (2014). How do we recognise who is speaking. *Frontiers In Bioscience*, *S6*(1), 92-109. doi: 10.2741/s417

Maurer, D., & Landis, T. (1990). Role of Bone Conduction in the Self-Perception of Speech. *Folia Phoniatrica Et Logopaedica*, *42*(5), 226-229. doi: 10.1159/000266070

McAleer, P., Todorov, A., & Belin, P. (2014). How Do You Say ‘Hello’? Personality Impressions from Brief Novel Voices. *Plos ONE*, *9*(3), e90779. doi: 10.1371/journal.pone.0090779

McGettigan, C. (2015). The social life of voices: studying the neural bases for the expression and perception of the self and others during spoken communication. *Frontiers in Human Neuroscience*, **9**(129), 1– 4. https://doi.org/10.3389/fnhum.2015.00129

McGettigan, C., Eisner, F., Agnew, Z.K., Manly, T., Wisbey, D., & Scott, S.K. (2013). T'ain't what you say, it's the way that you say it—left insula and inferior frontal cortex work in interaction with superior temporal regions to control the performance of vocal impersonations. *Journal of Cognitive Neuroscience*, **25**, 1875– 1886. https://doi.org/10.1162/jocn_a_00427

McGettigan, C., & Scott, S.K. (2014). Voluntary and involuntary processes affect the production of verbal and non-verbal signals by the human voice. *Behavioral and Brain Sciences*, **37**, 564– 565. https://doi.org/10.1017/S0140525X13004123

Meyer, M., & Lieberman, M. (2018). Why People Are Always Thinking about Themselves: Medial Prefrontal Cortex Activity during Rest Primes Self-referential Processing. *Journal Of Cognitive Neuroscience*, *30*(5), 714-721. doi: 10.1162/jocn_a_01232

Mezulis, A., Abramson, L., Hyde, J., & Hankin, B. (2004). Is There a Universal Positivity Bias in Attributions? A Meta-Analytic Review of Individual, Developmental, and Cultural Differences in the Self-Serving Attributional Bias. *Psychological Bulletin*, *130*(5), 711-747. doi: 10.1037/0033-2909.130.5.711

Mills, T., Bunnell, H., & Patel, R. (2014). Towards Personalized Speech Synthesis for Augmentative and Alternative Communication. *Augmentative And Alternative Communication*, *30*(3), 226-236. doi: 10.3109/07434618.2014.924026

Milne, A., Bianco, R., Poole, K., Zhao, S., Oxenham, A., Billig, A., & Chait, M. (2020). An online headphone screening test based on dichotic pitch. *Behavior Research Methods*, *53*(4), 1551-1562. doi: 10.3758/s13428-020-01514-0

Mitchell, J., Macrae, C., & Banaji, M. (2006). Dissociable Medial Prefrontal Contributions to Judgments of Similar and Dissimilar Others. *Neuron*, *50*(4), 655-663. doi: 10.1016/j.neuron.2006.03.040

Mittal, B. (2006). I, me, and mine—how products become consumers' extended selves. *Journal Of Consumer Behaviour*, *5*(6), 550-562. doi: 10.1002/cb.202

Mochon, D., Norton, M., & Ariely, D. (2012). Bolstering and restoring feelings of competence via the IKEA effect. *International Journal Of Research In Marketing*, *29*(4), 363-369. doi: 10.1016/j.ijresmar.2012.05.001

Montepare, J., Kempler, D., & McLaughlin-Volpe, T. (2014). The Voice of Wisdom. *Journal Of Language And Social Psychology*, *33*(3), 241-259. doi: 10.1177/0261927x13519080

Moore, J., & Obhi, S. (2012). Intentional binding and the sense of agency: A review. *Consciousness And Cognition*, *21*(1), 546-561. doi: 10.1016/j.concog.2011.12.002

Moran, J., Heatherton, T., & Kelley, W. (2009). Modulation of cortical midline structures by implicit and explicit self-relevance evaluation. *Social Neuroscience*, *4*(3), 197-211. doi: 10.1080/17470910802250519

Muth, F., Wirth, R., & Kunde, W. (2020). Temporal binding past the Libet clock: testing design factors for an auditory timer. *Behavior Research Methods*, *53*(3), 1322-1341. doi: 10.3758/s13428-020-01474-5

Nathanson, E. (2017). Native voice, self-concept and the moral case for personalized voice technology. *Disability and Rehabilitation*, **39**(1), 73– 81. https://doi.org/10.3109/09638288.2016.1139193

Nijhof, A., & Bird, G. (2019). Self-processing in individuals with autism spectrum disorder. *Autism Research*, *12*(11), 1580-1584. doi: 10.1002/aur.2200

Nijhof, A., Shapiro, K., Catmur, C., & Bird, G. (2020). No evidence for a common self-bias across cognitive domains. *Cognition*, *197*, 104186. doi: 10.1016/j.cognition.2020.104186

Northoff, G. (2013). Brain and self – a neurophilosophical account. *Child And Adolescent Psychiatry And Mental Health*, *7*(1), 28. doi: 10.1186/1753-2000-7-28

Northoff, G., & Hayes, D. (2011). Is Our Self Nothing but Reward?. *Biological Psychiatry*, *69*(11), 1019-1025. doi: 10.1016/j.biopsych.2010.12.014

Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, H., & Panksepp, J. (2006). Self-referential processing in our brain—A meta-analysis of imaging studies on the self. *Neuroimage*, *31*(1), 440-457. doi: 10.1016/j.neuroimage.2005.12.002

Norton, M., Mochon, D., & Ariely, D. (2012). The IKEA effect: When labor leads to love. *Journal Of Consumer Psychology*, *22*(3), 453-460. doi: 10.1016/j.jcps.2011.08.002

Nygren-Bonnier, M., Normi, L., Klefbeck, B., & Biguet, G. (2010). Experiences of decreased lung function in people with cervical spinal cord injury. *Disability And Rehabilitation*, *33*(6), 530-536. doi: 10.3109/09638288.2010.505995

Ohata, R., Asai, T., Imaizumi, S., & Imamizu, H. (2021). My voice therefore I spoke: sense of agency over speech enhanced in hearing self-voice. *BioRxiv*. https://doi.org/10.1101/2020.11.20.392308

Obhi, S., Swiderski, K., & Brubacher, S. (2012). Induced power changes the sense of agency. *Consciousness And Cognition*, *21*(3), 1547-1550. doi: 10.1016/j.concog.2012.06.008

Oyserman, D. (2001). Self-concept and identity. In A. Tesser & N. Schwarz, The Blackwell Handbook of Social Psychology (pp. 499-517). Malden, MA: Blackwell.

Oyserman, D., Elmore, K., & Smith, G. (2012). Self, self-concept, and identity. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (2nd ed., pp. 69–104). New York: Guilford Press.

Paladino, M., Mazzurega, M., Pavani, F., & Schubert, T. (2010). Synchronous Multisensory Stimulation Blurs Self-Other Boundaries. *Psychological Science*, *21*(9), 1202-1207. doi: 10.1177/0956797610379234

Pandey J. (2011) Source Memory. In: Kreutzer J.S., DeLuca J., Caplan B. (eds) Encyclopedia of Clinical Neuropsychology. Springer, New York, NY. https://doi.org/10.1007/978-0-387-79948-3_1159

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. J. Acoust. Soc. Am. 119, 2382–2393. doi: 10.1121/1.2178720

Pardo, J., Gibbons, R., Suppes, A., & Krauss, R. (2012). Phonetic convergence in college roommates. *Journal Of Phonetics*, *40*(1), 190-197. doi: 10.1016/j.wocn.2011.10.001

Pariyadath, V., & Eagleman, D. (2007). The Effect of Predictability on Subjective Duration. *Plos ONE*, *2*(11), e1264. doi: 10.1371/journal.pone.0001264

Patané, I., Brozzoli, C., Koun, E., Frassinetti, F., & Farnè, A. (2021). Me, you, and our object: Peripersonal space recruitment during executed and observed actions depends on object ownership. *Journal Of Experimental Psychology: General*, *150*(7), 1410-1422. doi: 10.1037/xge0001001

Payne, B., Lavan, N., Knight, S., & McGettigan, C. (2020). Perceptual prioritization of self-associated voices. *British Journal Of Psychology*, *112*(3), 585-610. doi: 10.1111/bjop.12479

Payne, S., Tsakiris, M., & Maister, L. (2017). Can the Self Become Another? Investigating the Effects of Self-Association with a New Facial Identity. *Quarterly Journal Of Experimental Psychology*, *70*(6), 1085-1097. doi: 10.1080/17470218.2015.1137329

Peng, Z., Wang, Y., Meng, L., Liu, H., & Hu, Z. (2019). One's own and similar voices are more attractive than other voices. *Australian Journal Of Psychology*, *71*(3), 212-222. doi: 10.1111/ajpy.12235

Pfister, R., Obhi, S., Rieger, M., & Wenke, D. (2014). Action and perception in social contexts: intentional binding for social action effects. *Frontiers In Human Neuroscience*, *8*. doi: 10.3389/fnhum.2014.00667

Pisanski, K., Oleszkiewicz, A., & Sorokowska, A. (2016). Can blind persons accurately assess body size from the voice?. *Biology Letters*, *12*(4), 20160063. doi: 10.1098/rsbl.2016.0063

Pohlert, T. (2014). The pairwise comparison of mean ranks package (PMCMR). R Package. Retrieved from https://CRAN.R-project.org/package=PMCMR.

Porcaro, C., Singer, C., Djokic, B., Danesh, A., Tappen, R., & Engstrom, G. (2021). Perceived Voice Disorders in Older Adults and Impact on Social Interactions. *Perspectives Of The ASHA Special Interest Groups*, *6*(1), 143-150. doi: 10.1044/2020_persp-20-00059

Pribuisiene, R., Uloza, V., Kupcinskas, L., & Jonaitis, L. (2006). Perceptual and Acoustic Characteristics of Voice Changes in Reflux Laryngitis Patients. *Journal Of Voice*, *20*(1), 128-136. doi: 10.1016/j.jvoice.2004.12.001

Pullin, G., Treviranus, J., Patel, R., & Higginbotham, J. (2017). Designing interaction, voice, and inclusion in AAC research. *Augmentative And Alternative Communication*, *33*(3), 139-148. doi: 10.1080/07434618.2017.1342690

Puts, D., Gaulin, S., & Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution And Human Behavior*, *27*(4), 283-296. doi: 10.1016/j.evolhumbehav.2005.11.003

Qin, P., & Northoff, G. (2011). How is our self related to midline regions and the default-mode network?. *Neuroimage*, *57*(3), 1221-1233. doi: 10.1016/j.neuroimage.2011.05.028

Quinlan, C., & Taylor, T. (2013). Enhancing the production effect in memory. *Memory*, *21*(8), 904-915. doi: 10.1080/09658211.2013.766754

R Core Team (2013). A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/.

Rabellino, D., Frewen, P., McKinnon, M., & Lanius, R. (2020). Peripersonal Space and Bodily Self-Consciousness: Implications for Psychological Trauma-Related Disorders. *Frontiers In Neuroscience*, *14*. doi: 10.3389/fnins.2020.586605

Radtke, T., Liszewska, N., Horodyska, K., Boberska, M., Schenkel, K., & Luszczynska, A. (2019). Cooking together: The IKEA effect on family vegetable intake. *British Journal Of Health Psychology*, *24*(4), 896-912. doi: 10.1111/bjhp.12385

Repp, B. (1987). The sound of two hands clapping: An exploratory study. *The Journal Of The Acoustical Society Of America*, *81*(4), 1100-1109. doi: 10.1121/1.394630

Repp, B., & Knoblich, G. (2004). Perceiving Action Identity. *Psychological Science*, *15*(9), 604-609. doi: 10.1111/j.0956-7976.2004.00727.x

Rosa, C., Lassonde, M., Pinard, C., Keenan, J., & Belin, P. (2008). Investigations of hemispheric specialization of self-voice recognition. *Brain And Cognition*, *68*(2), 204-214. doi: 10.1016/j.bandc.2008.04.007

Ruess, M., Thomaschke, R., & Kiesel, A. (2017). The time course of intentional binding. *Attention, Perception, & Psychophysics*, *79*(4), 1123-1131. doi: 10.3758/s13414-017-1292-y

Schäfer, S., Wentura, D., & Frings, C. (2017). Distinctiveness effects in self-prioritization. *Visual Cognition*, *25*(1-3), 399-411. doi: 10.1080/13506285.2017.1346739

Schäfer, S., Wentura, D., & Frings, C. (2020). Creating a network of importance: The particular effects of self-relevance on stimulus processing. *Attention, Perception, & Psychophysics*, *82*(7), 3750-3766. doi: 10.3758/s13414-020-02070-7

Schäfer, S., Wesslein, A., Spence, C., Wentura, D., & Frings, C. (2016). Self-prioritization in vision, audition, and touch. *Experimental Brain Research*, *234*(8), 2141-2150. doi: 10.1007/s00221-016-4616-6

Scherer, K. (1978). Personality inference from voice quality: The loud voice of extroversion. *European Journal Of Social Psychology*, *8*(4), 467-487. doi: 10.1002/ejsp.2420080405

Schroeder, J., & Epley, N. (2015). The Sound of Intellect. *Psychological Science*, *26*(6), 877-891. doi: 10.1177/0956797615572906

Schroeder, J., Kardas, M., & Epley, N. (2017). The Humanizing Voice: Speech Reveals, and Text Conceals, a More Thoughtful Mind in the Midst of Disagreement. *Psychological Science*, *28*(12), 1745-1762. doi: 10.1177/0956797617713798

Schroeder, S., Rembrandt, H., May, S., & Freeman, M. (2020). Does having a voice disorder hurt credibility?. *Journal Of Communication Disorders*, *87*, 106035. doi: 10.1016/j.jcomdis.2020.106035

Scott, S. K., and McGettigan, C. (2015). *The Voice APA Handbook of Nonverbal Communication*. Washington, DC: American Psychological Association.

Sell, A., Bryant, G., Cosmides, L., Tooby, J., Sznycer, D., & von Rueden, C., Krauss, A., & Gurven, M. (2010). Adaptations in humans for assessing physical strength from the voice. *Proceedings Of The Royal Society B: Biological Sciences*, *277*(1699), 3509-3518. doi: 10.1098/rspb.2010.0769

Serino, A., Bassolino, M., Farnè, A., & Làdavas, E. (2007). Extended Multisensory Space in Blind Cane Users. *Psychological Science*, *18*(7), 642-648. doi: 10.1111/j.1467-9280.2007.01952.x

Sforza, A., Bufalari, I., Haggard, P., & Aglioti, S. (2010). My face in yours: Visuo-tactile facial stimulation influences sense of identity. *Social Neuroscience*, *5*(2), 148-162. doi: 10.1080/17470910903205503

Shang, Z., Tao, T., & Wang, L. (2016). Don't Always Prefer My Chosen Objects: Low Level of Trait Autonomy and Autonomy Deprivation Decreases Mere Choice Effect. *Frontiers In Psychology*, *7*. doi: 10.3389/fpsyg.2016.00524

Shapiro, K., Caldwell, J., & Sorensen, R. (1997). Personal names and the attentional blink: A visual "cocktail party" effect. *Journal Of Experimental Psychology: Human Perception And Performance*, *23*(2), 504-514. doi: 10.1037/0096-1523.23.2.504

Sidtis, D., & Kreiman, J. (2011). In the Beginning Was the Familiar Voice: Personally Familiar Voices in the Evolutionary and Contemporary Biology of Communication. *Integrative Psychological And Behavioral Science*, *46*(2), 146-159. doi: 10.1007/s12124-011-9177-4

Smith, D., & Patterson, R. (2005). Perception of speaker size and sex of vowel sounds. *The Journal Of The Acoustical Society Of America*, *117*(4), 2374-2374. doi: 10.1121/1.4809396

Souza, P., Gehani, N., Wright, R., & McCloy, D. (2013). The Advantage of Knowing the Talker. *Journal Of The American Academy Of Audiology*, *24*(08), 689-700. doi: 10.3766/jaaa.24.8.6

Sparks, S., Cunningham, S., & Kritikos, A. (2016). Culture modulates implicit ownership-induced self-bias in memory. *Cognition*, *153*, 89-98. doi: 10.1016/j.cognition.2016.05.003

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, and Computers*, **31**, 137– 149. https://doi.org/10.3758/BF03207704

Stein, T., Siebold, A., & van Zoest, W. (2016). Testing the idea of privileged awareness of self-relevant information. *Journal Of Experimental Psychology: Human Perception And Performance*, *42*(3), 303-307. doi: 10.1037/xhp0000197

Stetson, C., Cui, X., Montague, P., & Eagleman, D. (2006). Motor-Sensory Recalibration Leads to an Illusory Reversal of Action and Sensation. *Neuron*, *51*(5), 651-659. doi: 10.1016/j.neuron.2006.08.006

Stolte, M., Humphreys, G., Yankouskaya, A., & Sui, J. (2017). Dissociating Biases towards the Self and Positive Emotion. *Quarterly Journal Of Experimental Psychology*, *70*(6), 1011-1022. doi: 10.1080/17470218.2015.1101477

Strahilevitz, M., & Loewenstein, G. (1998). The Effect of Ownership History on the Valuation of Objects. *Journal Of Consumer Research*, *25*(3), 276-289. doi: 10.1086/209539

Sui, J., He, X., & Humphreys, G. (2012). Perceptual effects of social salience: Evidence from self-prioritization effects on perceptual matching. *Journal Of Experimental Psychology: Human Perception And Performance*, *38*(5), 1105-1117. doi: 10.1037/a0029792

Sui, J., & Humphreys, G. (2015). The Interaction between Self-Bias and Reward: Evidence for Common and Distinct Processes. *Quarterly Journal Of Experimental Psychology*, *68*(10), 1952-1964. doi: 10.1080/17470218.2015.1023207

Sui, J., & Humphreys, G. (2017). Aging enhances cognitive biases to friends but not the self. *Psychonomic Bulletin & Review*, 24, 2021– 2030. https://doi.org/10.3758/s13423-017-1264-1

Sui, J., Rotshtein, P., & Humphreys, G. (2013). Coupling social attention to the self forms a network for personal significance. *Proceedings Of The National Academy Of Sciences*, *110*(19), 7607-7612. doi: 10.1073/pnas.1221862110

Suzuki, K., Lush, P., Seth, A., & Roseboom, W. (2019). Intentional Binding Without Intentional Action. *Psychological Science*, *30*(6), 842-853. doi: 10.1177/0956797619842191

Symons, C.S., & Johnson, B.T. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin*, **121**, 371– 394. https://doi.org/10.1037/0033-2909.121.3.371

Tacikowski, P., & Nowicka, A. (2010). Allocation of attention to self-name and self-face: An ERP study. *Biological Psychology*, *84*(2), 318-324. doi: 10.1016/j.biopsycho.2010.03.009

Tajadura-Jiménez, A., Grehl, S., & Tsakiris, M. (2012). The Other in Me: Interpersonal Multisensory Stimulation Changes the Mental Representation of the Self. *Plos ONE*, *7*(7), e40682. doi: 10.1371/journal.pone.0040682

Takahata, K., Takahashi, H., Maeda, T., Umeda, S., Suhara, T., Mimura, M., & Kato, M. (2012). It's Not My Fault: Postdictive Modulation of Intentional Binding by Monetary Gains and Losses. *Plos ONE*, *7*(12), e53421. doi: 10.1371/journal.pone.0053421

Truong, G., & Todd, R. (2017). SOAP Opera: Self as Object and Agent in Prioritizing Attention. *Journal Of Cognitive Neuroscience*, *29*(6), 937-952. doi: 10.1162/jocn_a_01083

Truong, G., Chapman, C., Chisholm, J., Enns, J., & Handy, T. (2016). Mine in motion: How physical actions impact the psychological sense of object ownership. *Journal Of Experimental Psychology: Human Perception And Performance*, *42*(3), 375-385. doi: 10.1037/xhp0000142

Tsakiris, M. (2008). Looking for Myself: Current Multisensory Input Alters Self-Face Recognition. *Plos ONE*, *3*(12), e4040. doi: 10.1371/journal.pone.0004040

Turk, D., Brady-van den Bos, M., Collard, P., Gillespie-Smith, K., Conway, M., & Cunningham, S. (2012). Divided attention selectively impairs memory for self-relevant information. *Memory & Cognition*, *41*(4), 503-510. doi: 10.3758/s13421-012-0279-0

Turk, D., Cunningham, S., & Macrae, C. (2008). Self-memory biases in explicit and incidental encoding of trait adjectives. *Consciousness And Cognition*, *17*(3), 1040-1045. doi: 10.1016/j.concog.2008.02.004

Turk, D., van Bussel, K., Waiter, G., & Macrae, C. (2011). Mine and Me: Exploring the Neural Basis of Object Ownership. *Journal Of Cognitive Neuroscience*, *23*(11), 3657-3668. doi: 10.1162/jocn_a_00042

Ulloa, J., Vastano, R., George, N., & Brass, M. (2019). The impact of eye contact on the sense of agency. *Consciousness And Cognition*, *74*, 102794. doi: 10.1016/j.concog.2019.102794

van den Bos, M., Cunningham, S., Conway, M., & Turk, D. (2010). Mine to remember: The impact of ownership on recollective experience. *Quarterly Journal Of Experimental Psychology*, *63*(6), 1065-1071. doi: 10.1080/17470211003770938

van der Meer, L., Costafreda, S., Aleman, A., & David, A. (2010). Self-reflection and the brain: A theoretical review and meta-analysis of neuroimaging studies with implications for schizophrenia. *Neuroscience & Biobehavioral Reviews*, *34*(6), 935-946. doi: 10.1016/j.neubiorev.2009.12.004

van Heuven, W., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-UK: A New and Improved Word Frequency Database for British English. *Quarterly Journal Of Experimental Psychology*, *67*(6), 1176-1190. doi: 10.1080/17470218.2013.850521

Wagner, H. L. (1993). On measuring performance in category judgment studies of nonverbal behavior.Journal of Nonverbal Behavior,17(1), 3-28

Wang, L., Qi, Y., Li, L., & Jia, F. (2021). A Combined Effect of Self and Reward: Relationship of Self- and Reward-Bias on Associative Learning. *Frontiers In Psychology*, *12*. doi: 10.3389/fpsyg.2021.647443

Ward, E., Jarman, L., Cornwell, P., & Amsters, D. (2016). Impact of voice and communication deficits for individuals with cervical spinal cord injury living in the community. *International Journal Of Language & Communication Disorders*, *51*(5), 568-580. doi: 10.1111/1460-6984.12232

Wickenden, M. (2011). Whose Voice is That?: Issues of Identity, Voice and Representation Arising in an Ethnographic Study of the Lives of Disabled Teenagers who use Augmentative and Alternative Communication (AAC). *Disability Studies Quarterly*, *31*(4). doi: 10.18061/dsq.v31i4.1724

Wilkinson, S., & Bell, V. (2016). The Representation of Agents in Auditory Verbal Hallucinations. *Mind & Language*, *31*(1), 104-126. doi: 10.1111/mila.12096

Wolf, W. (1931). Zuordnung individueller Gangmerkmale zur Individual-Charakteristik. Beihefte zur Zeitschrift für Angewandte Psychologie, 58, 108–122.

Wolpe, N., Haggard, P., Siebner, H., & Rowe, J. (2013). Cue integration and the perception of action in intentional binding. *Experimental Brain Research*, *229*(3), 467-474. doi: 10.1007/s00221-013-3419-2

Woods, K., Siegel, M., Traer, J., & McDermott, J. (2017). Headphone screening to facilitate web-based auditory experiments. *Attention, Perception, & Psychophysics*, *79*(7), 2064-2072. doi: 10.3758/s13414-017-1361-2

Woźniak, M., & Knoblich, G. (2019). Self-prioritization of fully unfamiliar stimuli. *Quarterly Journal Of Experimental Psychology*, *72*(8), 2110-2120. doi: 10.1177/1747021819832981

Xu, M., Homae, F., Hashimoto, R., & Hagiwara, H. (2013). Acoustic cues for the recognition of self-voice and other-voice. *Frontiers In Psychology*, *4*. doi: 10.3389/fpsyg.2013.00735

Yankouskaya, A., Humphreys, G., Stolte, M., Stokes, M., Moradi, Z., & Sui, J. (2017). An anterior–posterior axis within the ventromedial prefrontal cortex separates self and reward. *Social Cognitive And Affective Neuroscience*, *12*(12), 1859-1868. doi: 10.1093/scan/nsx112

Yin, S., Bi, T., Chen, A., & Egner, T. (2021). Ventromedial Prefrontal Cortex Drives the Prioritization of Self-Associated Stimuli in Working Memory. *The Journal Of Neuroscience*, *41*(9), 2012-2023. doi: 10.1523/jneurosci.1783-20.2020

Yin, X., Ma, Y., Xu, X., & Yang, H. (2018). The effect of self-referencing on memory for different kinds of source information. *Memory*, *27*(4), 519-527. doi: 10.1080/09658211.2018.1532009

Yoo, J., & Park, M. (2016). The effects of e-mass customization on consumer perceived value, satisfaction, and loyalty toward luxury brands. *Journal Of Business Research*, *69*(12), 5775-5784. doi: 10.1016/j.jbusres.2016.04.174

Zheng, Z., MacDonald, E., Munhall, K. and Johnsrude, I., 2011. Perceiving a Stranger's Voice as Being One's Own: A ‘Rubber Voice’ Illusion?. *PLoS ONE*, 6(4), p.e18655.

Zopf, R., Polito, V., & Moore, J. (2018). Revisiting the link between body and agency: visual movement congruency enhances intentional binding but is not body-specific. *Scientific Reports*, *8*(1). doi: 10.1038/s41598-017-18492-7

Appendix

Appendix 1. Full R model outputs for analyses:

	Estimates	*CI*
(Intercept)	581.29	555.40 – 607.18
Main Effect Voice Identity		
Voice identity (Friend)	53.57	34.67 – 72.48
Voice identity (Other)	62.22	43.19 – 81.24
Main Effect Trial Type		
Trial Type (Mismatch)	132.12	112.87 – 151.37
*Interaction Voice Identity * Trial Type*		
Voice identity (Friend) * Trial Type (Mismatch)	-45.34	-72.72 – -17.97
Voice identity (Other) * Trial Type (Mismatch)	-62.65	-90.18 – -35.12

Supplemental Table 1. Model output for the analysis of the reaction times in Experiment 1. The reference category for Voice Identity is "Self"; the reference category for Trial Type is "Match".

	Estimates	*CI*
(Intercept)	2.62	2.25 – 2.98
Main Effect Voice Identity		
Voice identity (Friend)	-0.33	-0.64 – -0.02
Voice identity (Other)	-0.36	-0.67 – -0.06

Supplemental Table *2*. Model output for the analysis of the *sensitivity* in Experiment 1. The reference category for Voice Identity is "Self".

	Odds Ratios	*CI*
(Intercept)	12.85	8.98 – 18.38
Main Effect Voice Identity		
Voice identity (Friend)	0.7	0.53 – 0.93
Voice identity (Other)	0.64	0.48 – 0.84
Main Effect Trial Type		
Trial Type (Mismatch)	0.46	0.35 – 0.60
*Interaction Voice Identity * Trial Type*		
Voice identity (Friend) * Trial Type (Mismatch)	1.25	0.87 – 1.80
Voice identity (Other) * Trial Type (Mismatch)	1.3	0.91 – 1.86

Supplemental Table 3. Model output for the analysis of accuracy in Experiment 1. The reference category for Voice Identity is "Self", the reference category for Trial Type is "Match".

	Estimates	*CI*
(Intercept)	540.52	514.21 – 566.83
Main Effect Voice Identity		
Voice identity (Friend)	35.35	21.58 – 49.12
Voice identity (Other)	52.2	38.30 – 66.10
Main Effect Gender Matching		
Gender Matching (Yes)	0.68	-36.55 – 37.91
*Interaction Voice Identity * Gender Matching*		
Voice identity (Friend) * Gender Matching (Yes)	-5.02	-24.49 – 14.46
Voice identity (Other) * Gender Matching (Yes)	-1.5	-21.29 – 18.28

Supplemental Table 4. Model output for the analysis of the reaction times in Experiment 2. The reference category for Voice Identity is "Self", the reference category for Gender Matching is "No".

	Estimates	*CI*
(Intercept)	2.95	2.66 – 3.25
Main Effect Voice Identity		
Voice identity (Friend)	-0.08	-0.33 – 0.16
Voice identity (Other)	-0.28	-0.52 – -0.03
Main Effect Gender Matching		
Gender Matching (Yes)	-0.07	-0.49 – 0.36
*Interaction Voice Identity * Gender Matching*		
Voice identity (Friend) * Gender Matching (Yes)	0.1	-0.25 – 0.45
Voice identity (Other) * Gender Matching (Yes)	-0.12	-0.47 – 0.22

Supplemental Table 5. Model output for the analysis of sensitivity in Experiment 2. The reference category for Voice Identity is "Self", the reference category for Gender-Matching is "No".

	Odds Ratio	*CI*
(Intercept)	9.2	6.52 – 12.98
Main Effect Voice Identity		
Voice identity (Friend)	1.11	0.90 – 1.36
Voice identity (Other)	1.08	0.88 – 1.33
Main Effect Gender Matching		
Gender Matching (Yes)	1	0.61 – 1.61
*Interaction Voice Identity * Gender Matching*		
Voice identity (Friend) * Gender Matching (Yes)	0.94	0.70 – 1.26
Voice identity (Other) * Gender Matching (Yes)	0.98	0.73 – 1.32

Supplemental Table 6. Model output for the analysis of accuracy in Experiment 2. The reference category for Voice Identity is "Self", the reference category for Gender Matching is "No".

	Estimates	*CI*
(Intercept)	540.87	520.53 – 561.22
Main Effect Voice Identity		
Voice identity (Friend)	32.84	23.07 – 42.61
Voice identity (Other)	51.44	41.52 – 61.37
Main Effect Choice		
Choice (Yes)	5.73	-23.03 – 34.50
*Interaction Voice Identity * Choice*		
Voice identity (Friend) * Choice (Yes)	36.93	23.10 – 50.76
Voice identity (Other) * Choice (Yes)	15.49	1.47 – 29.51

Supplemental Table *7*. Model output for the analysis of the reaction times in Experiment 3. The reference category for Voice Identity is "Self", the reference category for Choice is "No".

	Estimates	*CI*
(Intercept)	3.1	2.89 – 3.32
Main Effect Voice Identity		
Voice identity (Friend)	-0.27	-0.45 – -0.09
Voice identity (Other)	-0.44	-0.62 – -0.26
Main Effect Choice		
Choice (Yes)	-0.19	-0.50 – 0.12
*Interaction Voice Identity * Choice*		
Voice identity (Friend) * Choice (Yes)	0.24	-0.01 – 0.49
Voice identity (Other) * Choice (Yes)	0.11	-0.14 – 0.36

Supplemental Table 8. Model output for the analysis of sensitivity in Experiment 3. The reference category for Voice Identity is "Self", the reference category for Choice is "No".

	Odds Ratio	*CI*
(Intercept)	18.77	14.76 – 23.86
Main Effect Voice Identity		
Voice identity (Friend)	0.8	0.67 – 0.97
Voice identity (Other)	0.45	0.38 – 0.53
Main Effect Choice		
Choice (Yes)	1.13	0.81 – 1.59
*Interaction Voice Identity * Choice*		
Voice identity (Friend) * Choice (Yes)	0.76	0.59 – 0.99
Voice identity (Other) * Choice (Yes)	1	0.78 – 1.28

Supplemental Table 9. Model output for the analysis of accuracy in Experiment 3. The reference category for Voice Identity is "Self", the reference category for Choice is "No".

	Estimates	*CI*
(Intercept)	342.35	130.93 – 553.78
Main effect (Cue-tone duration)		
Tone duration (100ms)	33.21	-29.59 – 96.01
Tone duration (200ms)	98.05	35.24 – 160.86
Tone duration (400ms)	56.34	-6.45 – 119.14

Supplemental Table 10. Model output for the analysis of mean interval estimates in Experiment 4 Pilot. The reference category for Cue-interval duration is "50ms".

	Estimates	*CI*
(Intercept)	316.53	129.75 – 503.31
Main Effect Voice Identity		
Voice Identity (Other)	12.17	3.05 – 21.29
Main effect Agency		
Agency (Passive)	51.54	42.42 – 60.66

Supplemental Table 11. Model output for the analysis of mean interval estimates in Experiment 4. The reference category for Voice Identity is "Self", the reference category for Agency is "Active".

	Estimates	*CI*
(Intercept)	325.77	144.45 – 507.10
Main Effect Voice Identity		
Voice Identity (Other)	30.47	16.17 – 44.77
Main effect Agency		
Agency (Passive)	23.60	9.24 – 37.96
*Interaction Voice Identity * Agency*		
Voice Identity (Other) * Agency (Passive)	3.18	-17.14 – 23.50

Supplemental Table 12. Model output for the analysis of mean interval estimates in Experiment 5. The reference category for Voice Identity is "Self", the reference category for Agency is "Active".

	Estimates	CI
(Intercept)	506.51	488.14 – 524.88
Main Effect Voice Identity		
Voice identity (Friend)	41.46	28.54 – 54.39
Voice identity (Other)	62.48	49.40 – 75.55
Main Effect Condition		
Condition (Experimental)	-39.90	-53.08 – -26.72
*Interaction Voice Identity * Condition*		
Voice identity (Friend) * Condition (Experimental)	28.72	10.61 – 46.82
Voice identity (Other) * Condition (Experimental)	50.98	32.67 – 69.30

Supplemental Table 13. Model output for the analysis of reaction times in Experiment 6. The reference category for Voice Identity is "Self", the reference category for Condition is “Control” (i.e. new voice as self).

	Estimates	CI
(Intercept)	3.26	3.06 – 3.45
Main Effect Voice Identity		
Voice identity (Friend)	-0.12	-0.35 – 0.11
Voice identity (Other)	-0.37	-0.59 – -0.14
Main Effect Condition		
Condition (Experimental)	0.82	0.59 – 1.05
*Interaction Voice Identity * Condition*		
Voice identity (Friend) * Condition (Experimental)	-0.22	-0.55 – 0.10
Voice identity (Other) * Condition (Experimental)	-0.33	-0.66 – -0.01

Supplemental Table 14. Model output for the analysis of sensitivity in Experiment 6. The reference category for Voice Identity is "Self", the reference category for Condition is “Control” (i.e. new voice as self).

	Odds Ratio	CI
(Intercept)	18.13	14.03 – 23.45
Main Effect Voice Identity		
Voice identity (Friend)	0.88	0.68 – 1.15
Voice identity (Other)	0.57	0.44 – 0.73
Main Effect Condition		
Condition (Experimental)	4.16	2.71 – 6.38
*Interaction Voice Identity * Condition*		
Voice identity (Friend) * Condition (Experimental)	0.51	0.30 – 0.87
Voice identity (Other) * Condition (Experimental)	0.33	0.20 – 0.54

Supplemental Table 15. Model output for the analysis of accuracy in Experiment 6. The reference category for Voice Identity is "Self", the reference category for Condition is “Control” (i.e. new voice as self).

	Estimates	CI
(Intercept)	-73.90	-176.48 – 28.68
Main Effect Voice Identity		
Voice identity (Other)	19.45	6.75 – 32.15
Main Effect Agency		
Agency (Passive)	85.37	72.61 – 98.14
Main Effect Condition		
Condition (Experimental)	-8.89	-22.03 – 4.26
*Interaction Voice Identity * Agency*		
Voice identity (Other) * Agency (Passive)	-4.48	-30.11 – 21.14
*Interaction Condition * Agency*		
Condition (Experimental) * Agency (Passive)	-5.88	-24.04 – 12.29
*Interaction Voice * Condition*		
Voice identity (Other) * Condition (Experimental)	-2.22	-20.31 – 15.86

Supplemental Table 16. Model output for the analysis of mean interval estimates in Experiment 6. The reference category for Voice Identity is "Self", the reference category for Agency is "Passive", reference category for Condition is “Control” (i.e. new voice as self).

	Estimates	CI
(Intercept)	512.95	495.57 – 530.32
Main Effect Voice Identity		
Voice identity (Friend)	33.16	19.65 – 46.68
Voice identity (Other)	37.99	24.31 – 51.66
Main Effect Condition		
Condition (Experimental)	14.80	1.21 – 28.38
*Interaction Voice Identity * Condition*		
Voice identity (Friend) * Condition (Experimental)	-12.42	-31.58 – 6.73
Voice identity (Other) * Condition (Experimental)	-24.67	-44.05 – -5.30

Supplemental Table 17. Model output for the analysis of reaction times in Experiment 7. The reference category for Voice Identity is "Self", the reference category for Condition is “Control” i.e. the self-owned voice.

	Odds Ratio	*CI*
(Intercept)	25.53	19.65 – 33.16
Main Effect Voice Identity		
Voice identity (Friend)	1.03	0.73 – 1.44
Voice identity (Other)	0.46	0.35 – 0.62
Main Effect Condition		
Condition (Experimental)	0.64	0.47 – 0.87
*Interaction Voice Identity * Condition*		
Voice identity (Friend) * Condition (Experimental)	1.42	0.90 – 2.23
Voice identity (Other) * Condition (Experimental)	1.59	1.08 – 2.35

Supplemental Table 18. Model output for the analysis of accuracy in Experiment 7. The reference category for Voice Identity is "Self", the reference category for Condition is “Control” i.e. the self-owned voice.

	Estimates	*CI*
(Intercept)	282.47	98.62 – 466.31
Main Effect Voice Identity		
Voice identity (Other)	55.21	41.86 – 68.57
Main Effect Agency		
Agency (Passive)	89.09	75.77 – 102.41
Main Effect Condition		
Condition (Experimental)	13.76	0.40 – 27.13
*Interaction Voice Identity * Agency*		
Voice identity (Other) * Agency (Passive)	10.76	-8.09 – 29.62
*Interaction Condition * Agency*		
Condition (Experimental) * Agency (Passive)	-1.77	-20.66 – 17.13
*Interaction Voice * Condition*		
Voice identity (Other) * Condition (Experimental)	3.86	-15.02 – 22.75
*Interaction Voice Identity * Agency * Condition*		
Voice identity (Other) * Condition (Experimental) * Agency (Passive)	0.12	-26.57 – 26.81

Supplemental Table 19. Model output for the analysis of mean interval estimates in Experiment 7. The reference category for Voice Identity is "Self", the reference category for Agency is "Passive", the reference category for condition is “Control” i.e. the self-owned voice.

	Estimates	CI
(Intercept)	545.51	514.57 – 576.44
Main Effect Voice Identity		
Voice identity (Friend)	68.67	55.02 – 82.31
Voice identity (Other)	97.40	83.67 – 111.13
Main Effect Condition		
Condition (Experimental)	-21.83	-65.59 – 21.93
*Interaction Voice Identity * Condition*		
Voice identity (Friend) * Condition (Experimental)	-2.66	-21.91 – 16.59
Voice identity (Other) * Condition (Experimental)	-4.32	-23.74 – 15.10

Supplemental Table 20. Model output for the analysis of reaction times in Experiment 10. The reference category for Voice Identity is "Self", the reference category for Condition is “Control” i.e. the self-owned voice.

	Estimates	CI
(Intercept)	3.48	3.24 – 3.73
Main Effect Voice Identity		
Voice identity (Friend)	-0.17	-0.39 – 0.05
Voice identity (Other)	-0.34	-0.56 – -0.13
Main Effect Condition		
Condition (Experimental)	-0.01	-0.36 – 0.34
*Interaction Voice Identity * Condition*		
Voice identity (Friend) * Condition (Experimental)	0.11	-0.20 – 0.42
Voice identity (Other) * Condition (Experimental)	0.18	-0.13 – 0.49

Supplemental Table 21. Model output for the analysis of sensitivity in Experiment 10. The reference category for Voice Identity is "Self", the reference category for Condition is “Control” i.e. the self-owned voice.

	Odds Ratio	CI
(Intercept)	34.47	24.06 – 49.37
Main Effect Voice Identity		
Voice identity (Friend)	0.54	0.39 – 0.76
Voice identity (Other)	0.39	0.29 – 0.54
Main Effect Condition		
Condition (Experimental)	0.91	0.55 – 1.50
*Interaction Voice Identity * Condition*		
Voice identity (Friend) * Condition (Experimental)	1.70	1.05 – 2.74
Voice identity (Other) * Condition (Experimental)	1.19	0.76 – 1.85

Supplemental Table 22. Model output for the analysis of accuracy in Experiment 10. The reference category for Voice Identity is "Self", the reference category for Condition is “Control” i.e. the self-owned voice.

Appendix 2. Full stimulus set for the shopping task used in Experiment 8. Food and household-related items (216) were divided into three sets so that each set was broadly matched on word frequency according to the SUBTLE-X corpus, syllable length and item length.

SET 1		SET 2		SET 3	
apples	oven rack	anchovies	hoover	apricots	mangos
apron	paper	asparagus	kiwis	artichokes	marrow
aubergine	peaches	beetroot	lawnmower	baby food	matches
basil	peeler	biscuits	lemonade	bacon	meatballs
blender	pitta bread	blanket	lemons	bananas	melon
cereal	plates	blueberries	maple syrup	bathmat	meringue
cheese	pomegranate	bowls	moisturiser	batteries	milk
chilli	porridge	brownie	mushrooms	bleach	olive oil
coconut	pudding	cashews	mussels	bread	olives
cod	raisins	casserole dish	oatcakes	broccoli	torch
conditioner	raspberries	celery	oranges	butter	pasta
coriander	rice	chard	parsley	cabbage	pears
crumpets	salt	cheesecake	pastry	carrots	pepperoni
cucumber	sandwich	chicken	peanuts	cherries	perfume
dog food	scones	chives	peppers	chewing gum	pickles
electric whisk	shallots	chutney	pineapples	chocolate	plunger
extension lead	shampoo	cinnamon	pizza	courgettes	popcorn
gammon	spinach	coasters	plant	custard	potatoes
granola	spoon	coffee	plums	dusters	pretzel
halibut	squash	cookies	sage	feta	prunes
honey	steak	cordial	salad cream	fridge	pumpkins
kale	steamer	crab	salsa	ginger	quiche
lentils	strainer	crackers	sausages	guacamole	rhubarb
lettuce	strimmer	cream	scales	haddock	saffron
limes	sugar	crisps	sieve	horseradish	scented candle
lollipops	swede	cushion	soap	hummus	scissors
marmalade	sweetcorn	eggs	spaghetti	icing sugar	shrimp
marmite	tangerines	envelopes	stapler	jam	soft toy
mayonnaise	teapot	flannel	string	juicer	sponges
microwave	toaster	freezer	strudel	ketchup	tofu
mint	tomatoes	garlic	sun-tan lotion	kettle	trout
muffins	towels	grapes	sweetener	knife	tumble dryer
mugs	turmeric	grater	trifle	lamp	wasabi
mustard	turnips	hair dye	tuna	lasagne	ladle
nappies	walnuts	ham	turkey	Laundry powder	whiskey
onions	yoghurt	Hand wash	vinegar	lobster	wine

Appendix 3. Cue stimuli used in Experiments 6 and 7 to elicit the target word, “hello”, in participants own voices.

Trial	Utterance
Practice Trial	"Hello, how are you today?"
Trial 1	"I just called to say hello"
Trial 2	"Hello"
Trial 3	"Oh, hello, could I speak to Dr Johns please?"
Trial 4	"Hello, can you hear me?"
Trial 5	"Hello"
Trial 6	"Hello, are you there?"
Trial 7	"Hello, is that Alex?"
Trial 8	"Hello"
Trial 9	"Hello, how are you today?"
Trial 10	"Hello…hello…. hello…. hello…hello…"(said in a cheerful voice)
Trial 11	"Hello…hello…. hello…. hello…hello…"(said in a neutral voice)
Trial 12	"Hello…hello…. hello…. hello…hello…"(said in a sad voice)