CHORAL SINGING AND EMOTIONAL COMMUNICATION

IN PEOPLE WITH PARKINSON'S DISEASE

by

Esztella Vezer

Honours Bachelor of Arts, Ryerson University, 2014

A thesis

presented to Ryerson University

in partial fulfillment of the

requirements for the degree of

Master of Arts

in the program of

Psychology

Toronto, Ontario, Canada, 2016

AUTHOR'S DECLARATION FOR ELECTRONIC SUBMISSION OF A THESIS

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my thesis may be made electronically available to the public.

Choral Singing and Emotional Communication

in People with Parkinson's Disease

Esztella Vezer

Master of Arts, 2016

Psychology

Ryerson University

Abstract

A large part of successful interpersonal communication relies on both the production and interpretation of vocal and facial expressions. Both of these abilities appear to be impaired in people with Parkinson's disease (PD), a neurodegenerative disease affecting the muscles of the body. Based on positive effects of rhythmic auditory stimulation on overall motor functioning and the effects of intentional simulation of expression on facial mimicry, the present study investigated the impact of a choir program involving singing of happy and sad songs on a group of 12 PD patients. Participants were tested before and after program completion on measures of: (1) facial mimicry of a range of emotional expressions; (2) emotion perception, and; (3) vocal quality. It was found that choir participation improved some measures of vocal quality, while effects on facial mimicry and emotion perception were marginal.

Acknowledgments

I would like to extend my sincerest gratitude to my supervisor, Dr. Frank Russo, not only for his support and guidance throughout this project, but for inspiring me to become involved in this area of research more than four years ago. The Science of Music, Auditory Research and Technology (SMART) Lab has been my academic home since my undergraduate studies and I cannot imagine a more nurturing environment in which to conduct research. I would also like to thank my supervisory committee member Dr. Todd Girard, whose knowledge and attention to detail have undoubtedly improved the quality of this thesis. Gratitude is also owed to Dr. Alexandra Fiocco for generously agreeing to invest her time and insights as a reader on this project.

The help of several current and former members of SMART Lab was also instrumental to this thesis. Among the research assistants who volunteered their time to assist with a great many tasks, Domenica Fanelli deserves a special thanks for her extensive commitment and passion for this project. My gratitude also extends to former SMART Lab postdoctoral fellow Dr. Steven Livingstone, who was able to provide crucial guidance despite geographic constraints and time limitations.

A tremendous amount of thanks is also owed to the organizations who have made the Parkinson's choir possible, from Parkinson Society Canada for help with recruitment to the Royal Conservatory of Music for providing a beautiful setting for the choir. I am also beyond grateful to the people who breathed life into the choir, which not only include the highly talented and much-loved choir director Paula Wolfson and pianist Sina Fallah, but also the participants themselves, without whom this research would not have been possible. I am not only grateful for their willingness to commit so much time and energy to this project, but am also inspired by the

way in which each individual contributed to building what turned out to be such an important community. Working with this community has made this journey immeasurably more rewarding, for which I am also grateful.

Finally, but no less importantly, I would like to thank my husband, family and friends (as well as my two cats) for their incredible patience and ongoing support throughout this endeavour. Though less conventional in her support, I also thank my baby daughter Matilda – due in December 2016 – whose kicks and wriggles punctuated my writing sessions and reminded me of the importance of taking breaks.

Table of Contents

List of Tables

List of Figures

Choral Singing and Emotional Communication

in People with Parkinson's Disease

Introduction

As social beings, humans rely heavily on interpersonal communication in order to lead adaptive lives (Ekman, 1992). With emotion being such a large part of our species' conscious repertoire, accurate interpretation of emotion in others is essential to our navigation of social relationships. We need to know when our partners are sad, our children are fearful or a stranger on the street is angry, in order to respond appropriately. Among the cues that we use to interpret emotion in others, facial and vocal expressions are a very important source of information; a silent smile, frown or tone of voice can convey a thousand words. In this respect, our faces and voices are like windows into our internal world, and when others with whom we are interacting are able to perceive and understand our internal world, successful communication and mutual understanding are facilitated.

Of course, the ability to produce emotional facial and vocal expressions is as important to successful communication as interpretation of emotion in others. If one cannot produce expressions consistent with his/her affective state, there is little hope for effective emotional communication with others. Similarly, unusually weakened emotional expressiveness that does not adequately convey the strength (or degree) of one's inner state can be confusing to the interpreter; intensity of an emotional facial or vocal expression is as informative as the expressed emotion itself, representing the potential seriousness of the triggering stimulus and the degree to which the person may have been affected.

Both the production and perception of emotional expressions are impaired in PD, a neurodegenerative disorder that affects approximately1% of individuals over the age of 60 (de

1

Lau & Breteler, 2006) and is related to the death of dopamine-producing cells in the substantia

nigra (part of the basal ganglia) that results in motor symptoms such as muscle rigidity, impaired

balance and gait, tremors and slowness of movement (Gelb, Oliver, & Gilman, 1999). The

disease may also result in mild cognitive impairments, in particular problems with executive

functioning (Dubois, Boller, Pillon, & Agid, 1991; Pillon et al., 1989), as well as emotional

deficits such as depression (Cummings, 1992), apathy (Pluck & Brown, 2002) and anxiety

(Shulman, Taback, Rabinstein, & Weiner, 2002).

Hypomimia, or facial muscle rigidity, is regarded as a secondary symptom of PD and

results in the characteristic lack of facial expressiveness known as the "masked face syndrome"

(Colcher & Simuni, 1999). Similarly, impairment of vocal muscles results in dysarthria, or

unclear or inarticulate speech, which also limits a person's capacity for vocal expressiveness

(Colcher & Simuni, 1999). Furthermore, impairments in perceiving emotion in others may be

related to the deficits in emotional expressiveness via the mechanism of spontaneous mimicry,

which refers to the automatic mirroring of the body language, movements, vocal and facial

expressions of the person with whom one is communicating in a face-to-face context. It is

believed that this mechanism potentially aids the understanding of others' emotions by way of

muscular feedback (Buck, 1980) and, when impaired, can inhibit such an understanding

(Niedenthal, Brauer, Halberstadt, & Innes-Ker, 2001; Oberman, Winkielman, & Ramachandran,

2007). If this theory is valid, then the deficit in emotion perception in PD may be due to a deficit

in the mimicry of facial expressions. Indeed, facial mimicry has recently been found to be

impaired in PD (Argaud et al., 2016; Livingstone, Vezer, McGarry, Lang, & Russo, 2016). It is

this set of impairments in emotional communication that is the target of the interventions

investigated in the present study, and will be described in greater detail in the sections that follow.

Although deficits in emotional communication in PD have received considerably less research attention than the primary symptoms of the disease, they nonetheless have a substantial effect on PD patients' lives, and are therefore in need of further study. Because production and recognition of emotion in facial and vocal expressions are so integral to human communication, deficits in these areas can have a detrimental impact on interpersonal relationships. For example, lack of expressiveness can give a false impression of withdrawal, lack of interest and coldness during interpersonal encounters (Pentland et al., 1987), and when PD patients attempt to smile, observers may perceive the smiles as insincere (Pitcairn, 1990). In contrast, smiling in neurotypicals is generally viewed as being reflective of emotional stability, agreeableness and sociability (Frank, Ekman, & Friesen, 1993).

This overall lack of emotional expressiveness can lead to frequent misunderstandings of how a person with PD is truly feeling (e.g., Ellgring et al., 1993; Macht, Schwarz, & Ellgring, 1999). Indeed, relationships with caregivers appear to be adversely affected by the disease (Aarsland, Larsen, Karlsen, Lim, & Tandberg, 1999; Martínez-Martín et al., 2005), and both PD patients and their caregivers have been found to experience diminished quality of life (Muslimovic, Post, Speelman, Schmand, & Haan, 2008; Schrag, Jahanshahi, & Quinn, 2000). The inability to express one's emotional state and interpret that of others may also eventually result in disengagement and isolation (Smith et al., 1996; Dujardin et al., 2004), contributing to the increased rates of depression and anxiety described earlier (Cummings, 1992; Pluck & Brown, 2002; Shulman, Taback, Rabinstein, & Weiner, 2002).

Production of Emotional Facial Expressions in PD

A number of researchers have specifically explored PD patients' ability to express emotion via facial expressions. In one such study, Simons, Pasqualini, Reddy, and Wood (2004) presented participants with PD and their spouses, who acted as healthy controls, with three separate tasks that were designed to elicit either voluntary or involuntary emotional facial expressions: (1) the viewing of amusing videos (involuntary expressiveness); (2) involvement in a pleasant conversation (involuntary expressiveness), and; (3) deliberate, trained production of emotional facial expressions that are incongruent with emotions expected to be elicited by amusing videos that were being watched concurrently (voluntary expressiveness). It was found that the PD group exhibited significantly lower levels of overall facial expressiveness than the control group, as measured by trained raters, with the greatest difference seen in involuntary expressiveness.

These findings replicated those of an earlier and similar study by Smith, Smith, and Ellgring (1996) in which the PD group demonstrated greater impairments in the involuntary production of facial expressions than the control group during a task that involved observing videos of emotional facial expressions. Conversely, comparably small impairments were found in the voluntary production of emotional facial expressions. Other studies have also confirmed a greater level of impairment in involuntary production of emotional facial expressions versus voluntary production (e.g., Simons, Ellgring, & Pasqualini, 2003). Nonetheless, the evidence discussed suggests that deficits in the latter do exist (e.g., Jacobs et al., 1995; Madeley et al., 1995). The greater impairment in involuntary versus voluntary expressions may be due to the distinct neural networks that control each. Specifically, voluntary expressions are initiated via cortical routes while involuntary expressions are initiated by subcortical routes (Rinn, 1984).

4

Because PD is a dysfunction of the basal ganglia – a subcortical structure – it is not surprising that greater impairments exist in involuntary expressiveness.

These findings suggest that people with PD likely experience most difficulty with emotional communication in everyday interpersonal contexts where expressiveness is expected to be elicited naturally rather than deliberately. However, given that voluntary facial expressiveness is relatively spared in terms of a deficit, it may be possible to leverage this ability via repeated practice in potential interventions. A study by Berenbaum & Rotter (1993) found that intensity of voluntary expressions were positively correlated with intensity of spontaneous expressions in neurotypicals, offering support that the two circuits overlap despite being distinct. Therefore, as well as strengthening the facial musculature, voluntary expressiveness exercises may provide an alternative route to strengthening connections in the motor cortex via a relatively intact motor pathway.

Perception of Emotion in Facial Expressions in PD

More recently, research has shown that impairments also exist in the ability of people with PD to perceive emotion in others' facial expressions (Assogna et al., 2010; Buxton et al., 2013; Dujardin et al., 2004; Gray & Tickle-Degnen, 2010; Suzuki, Hoshino, Shigemasu, & Kawamura, 2006). The nature and extent of impairment in this area are influenced by various factors, including type of emotion, intensity of perceived emotion, disease severity, and medication use, which are discussed independently in the following sections.

Type of emotion. The type of emotion expressed appears to affect accuracy levels. For example, perception of fear and disgust have been found by a number of researchers to be most impaired in PD (e.g., Assogna et al., 2010; Kan, Kawamura, Hasegawa, Mochizuki, & Nakamura, 2002; Suzuki et al., 2006), with negative emotions in general being harder to decode,

though findings across studies are not entirely consistent (Gray & Tickle-Degnen, 2010; Sprengelmeyer et al., 2003). Conversely, considerable evidence suggests that happiness is the most easily perceived emotion (Suzuki et al., 2006) with the least difference in recognition ability between healthy controls and PD samples (Dujardin et al., 2004). This is likely because happy expressions contain the least overlap with other expressions in terms of the required muscle movements, leading to easier identification relative to other emotions (Calvo & Lundquist, 2008).

In order to place these selective deficits into appropriate context, it is important to consider the profile of emotion recognition ability in healthy populations. There is ample evidence to suggest, for example, that fear is the most difficult emotion to recognize even among healthy populations (e.g., Biehl, 1997; Russell, 1994). The selective nature of emotion recognition ability also appears to change across the lifespan, with Calder and colleagues (2003) finding that older adults experience increasing difficulty identifying fear and, though somewhat less challenging, anger and sadness. Interestingly, older adults in this study displayed a slight advantage over young adults in the identification of disgust.

Cognition is of course integral to emotion recognition, therefore it is expected that normal age-related cognitive changes might be responsible for impairments in this ability across the lifespan. Furthermore, emotions that are more challenging to recognize even in the absence of cognitive impairment may become increasingly more difficult to recognize when cognitive problems are present. For example, the general difficulty in recognizing fearful expressions in neurotypicals may be even more pronounced in individuals with cognitive impairment, owing to experiencing difficult cognitive tasks as a lot more challenging. Indeed, researchers have established an explicit link between even mild cognitive impairment and problems with emotion

perception (McCade, Savage, & Naismith, 2012). Since PD is a condition that predominantly presents in older adulthood and whose symptom profile often includes cognitive deficits, cognitive factors most likely play a role in impaired emotion recognition ability.

However, these impairments may be the result of more than just cognitive deficits causing more difficulty with difficult tasks (as in the case of recognizing fear). Suzuki et al. (2006) pointed out that the impact of the specific neural pathology underlying PD may also explain impairments in the recognition of fear, as well as disgust. Specifically, the basal ganglia, which is a major brain structure affected in PD, appears to play a role in the recognition of disgust (e.g., Keane, Antoun, Manes, Calder, & Young, 2000; Winston, O'Doherty, & Dolan, 2003). Similarly, the amygdala has been increasingly associated with the ability to recognize expressions of fear (Calder, Lawrence, & Young, 2001; Whalen et al., 1998), and is another structure that appears to be impaired in PD (Harding, Stimson, Henderson, & Halliday, 2002; Sato et al., 2002).

Furthermore, neural pathways within the basal ganglia are also notably involved in motor systems, suggesting an additional and somewhat different potential link to emotion recognition ability that is unique to PD. It is this motor pathway – in the specific context of facial (and vocal) expressiveness – and its relationship to emotion perception that is the focus of the present study and will be discussed explicitly in a later section.

Intensity of perceived emotion. Intensity of the emotions being conveyed by facial expressions is also important to perception, as expressions of lower intensity, i.e., weaker expressions, are more difficult to perceive (Hess, Blairy, & Kleck, 1997). A handful of studies have used stimuli that vary in intensity, and therefore also in difficulty in terms of interpreting the target emotion. The use of weaker stimuli helps identify impairments that are apparent only

in response to subtle displays of emotional expressiveness. A common way of achieving varying levels of intensity and difficulty is to artificially increase or decrease the level of expressed emotion in an image by gradual increments with the use of computer software, and then to determine the percentage of healthy controls that accurately identify the target emotion for each intensity level (e.g., Buxton et al., 2013; Dujardin et al., 2004). To obtain varying task difficulty levels, Buxton and colleagues (2013) selected three levels of intensity for their stimuli based on accuracy percentages achieved by a control group on an emotion identification task involving seven possible emotions (difficult = 50% accuracy, moderate = 70% accuracy, easy = 100% accuracy), while Dujardin et al. (2004) used two levels (30% and 70%).

As discussed previously, all of these studies demonstrated deficits in emotion perception in the PD group, and the accuracy differences between the PD groups and control groups were increasingly more pronounced in greater task difficulty conditions (Buxton et al., 2013; Dujardin et al., 2004), most notably in the case of happy expressions, the recognition of which had been considered relatively intact by studies using less difficult stimuli (Buxton et al., 2013). This finding suggests that tasks consisting of more subtle (or difficult) expressions possess the most optimal discriminating power to reveal the true extent of emotion perception deficits in PD samples.

Disease severity and medications. The PD sample used by Suzuki et al. (2006), who found impairments in the perception of disgust only, consisted almost entirely of medicated PD patients at an early stage of disease progression, raising the question of how disease severity and medication impact performance on emotion recognition tasks. At these early stages of disease progression, it could be that impairments are less pronounced, especially in medicated PD samples such as Suzuki and colleagues' (2006). Using the Penn Emotional Recognition Test, a

database consisting of three-dimensional digital still images of faces depicting either felt or evoked emotional facial expressions, Assogna et al. (2010) obtained results similar to those of Suzuki et al. (2006) in a medicated PD sample at a moderate stage of disease progression.

Interestingly, while unmedicated PD participants in early stages of the disease demonstrated impairments in the perception of a broader range of emotional facial expressions, Sprengelmeyer et al. (2003) found that their medicated but more severely affected PD sample did not show the same deficits. Similarly, Lawrence, Goerendt, and Brooks (2007) showed that temporary withdrawal from anti-Parkinson's medication resulted in impairments in the perception of angry facial expressions. It is likely that medication is a potentially mitigating factor that might limit the extent of impairment in the recognition of certain emotions.

Spontaneous Mimicry

The ability to produce and interpret emotion via facial expressions has recently been linked to spontaneous facial mimicry, which is part of a larger phenomenon known as emotional contagion (Hatfield, Cacioppo, & Rapson, 1993), or simply spontaneous mimicry (Lundqvist & Dimberg, 1995). These terms refer to the moment-to-moment mirroring of behaviours between two interacting individuals; a person appears to unwittingly take on the speech patterns and tones (Goldinger, 1998; Neumann & Strack, 2000), facial expressions (Dimberg & Öhman, 1996; Dimberg, Thunberg, & Elmehed, 2000), gestures and postures (Chartrand & Bargh, 1999) of the person with whom they are communicating.

Facial mimicry, then, specifically refers to the spontaneous, involuntary imitation of others' facial expressions during face-to-face communication. It is worth noting that spontaneous mimicry applies in the same way to vocal expressiveness, with auditory rather than visual observation triggering mimicry activity of the vocal muscles. However, investigation of

spontaneous mimicry in emotion research has tended to focus on spontaneous *facial* mimicry, most likely due to the less invasive nature of facial muscle tracking.

The facial feedback hypothesis, an embodied theory of emotional communication, puts forth that spontaneous mimicry functions as an integral mechanism in emotion understanding (Buck, 1980). Though there are divided schools of thought in terms of the precise nature of this mechanism (see Hatfield et al., 1993; Laird et al., 1994; Lundqvist & Dimberg, 1995), there is consensus with respect to its role in facilitating accurate perception of emotion in others, the evidence for which will be discussed in greater detail in a later section.

It is important to highlight that facial mimicry differs from the kind of involuntary production of emotional facial expressions referred to in earlier discussions of works by Simons et al. (2004) and Smith et al. (1996), in that mimicry is not necessarily directly observable and is more rapid in its presentation. Accordingly, the accepted method of measuring mimicry – electromyography (EMG) – affords a level of sensitivity far superior to what can be observed by the naked eye, both in terms of identifying temporal onset and the strength of muscle activity. Though directly observable spontaneous facial expressions may relate to mimicry, the latter exclusively refers to discreet facial muscle movements that are optimally measured using EMG.

Facial mimicry in response to emotional facial expressions. Initial support for the presence of an innate, predictable mimicry response during observation of emotional facial stimuli came from studies of healthy populations. Mimicry research has shown that presentation of positive and negative facial expressions trigger patterns of involuntary activity in facial muscles. Specifically, happy facial expressions evoke increased activity in the zygomaticus major muscle (used for smiling), while increased activity in the corrugator supercilii and the medial frontalis muscles (used for frowning) is seen in response to angry, sad and fearful facial

expressions (Dimberg et al., 2000; Lundqvist, 1995). This response is typically found to take place within 300 ms following presentation of the stimuli (Dimberg & Thunberg, 1998).

Other research has extended these findings to dynamic audio-visual presentations of emotional song, concluding that these also elicit facial mimicry in observers. Livingstone, Thompson, and Russo (2009) recorded facial muscle movements via EMG while asking participants to observe a video clip of a target singing a line of song and to repeat the sung line several seconds following presentation of the stimuli. Spontaneous facial mimicry was detected at onset of the target, even prior to participants beginning to sing, though this could have been interpreted as evidence of motor preparation rather than a response to perceived emotion alone.

Chan, Livingstone, and Russo (2013) sought clarification in this regard with a study that required participants to either passively observe videos of targets singing with various emotions (happy, sad and neutral) or imagine singing the song directly after its presentation. As before, facial muscle activity was recorded by EMG throughout the task. Spontaneous facial mimicry in the appropriate muscle regions at point of stimulus presentation was found in both conditions and to the same extent, suggesting that mimicry is not contingent on motor planning. Assuming that the imaginal group was adequately complying with instructions, the lack of difference in mimicry levels between the two groups also suggests that mimicry is an automatic response to emotion perception alone.

Facial mimicry and unconscious exposure to emotional facial expressions. Indeed, it appears that not only is observation of emotional expression enough to elicit a mimicry response, but the same remains true even when emotional stimuli are not consciously perceived. In a pivotal study, Dimberg, Thunberg, and Elmehed (2000) presented pictures of happy, angry and neutral emotional facial expressions taken from the Pictures of Facial Affect database (Ekman &

Friesen, 1976) to participants for durations of 30ms, which has widely been shown to fall below the threshold of conscious detection (e.g., Dimberg & Öhman, 1996; Morris et al., 1998; Whalen et al., 1998). Between the unconscious presentations of the images, neutral facial expressions were displayed for longer durations in order to mask the target stimuli. As with the previous studies, facial muscle activity in the zygomaticus and corrugator muscle regions was recorded by EMG throughout the experiment. The researchers found that unconscious presentation of happy faces triggered an increase in smiling muscle activity and decrease in frowning muscle activity, while angry faces triggered an increase in frowning activity and decrease in smiling muscle activity. The neutral unconscious condition triggered an intermediate level of activity in both muscle regions.

 Facial mimicry and impaired emotion perception. Returning to the facial feedback hypothesis (Buck, 1980), perhaps the most convincing work in the area of mimicry that supports a causal link between spontaneous muscle activity and perception of emotion is that of Niedenthal et al. (2001) and Oberman et al. (2007), who showed that mimicry may function to enhance the ability to perceive emotional facial expressions in others. Niedenthal et al. (2001) asked neurotypical participants to identify the point at which a happy face morphed into a sad face, and vice versa, while facial muscle activity was recorded by EMG. The novel aspect of this study was a condition in which participants held a pen horizontally between their teeth throughout the task, thus suppressing normal activity in the zygomaticus muscle region around the mouth. Participants in the pen-holding condition took a significantly longer time to identify a change between the two emotions, suggesting impairment in perceptual abilities that directly relate to inhibition of mimicry.

 In a similar experiment, Oberman et al. (2007), recorded five areas of facial muscle

activity by EMG while participants viewed images of faces with expressions of happiness, sadness, fear and disgust, and were asked to identify the observed emotions. Adapting the procedure used by Niedenthal and colleagues (2001), participants either clenched a pen between their teeth without letting it touch their lips (strong mimicry inhibition), held it loosely in their lips without letting it touch their teeth (no mimicry inhibition) or chewed gum (intermittent mimicry inhibition) during the tasks. It was found that the inhibited groups showed greater impairments in emotion perception, particularly the pen-biting group and particularly when responding to happy facial expressions. The authors suggested that the latter finding is due to the need for a greater number of muscles for the production of happy expressions, the majority of which were consistently occupied in the mimicry inhibition condition (Oberman et al., 2007).

In another novel study, researchers compared 16 women at a cosmetic clinic who underwent injections of botulinum toxin (Botox) to 15 women who underwent injections of Restylane, a dermal filler, on an emotion recognition task two weeks after receiving the cosmetic treatments (Neal & Chartrand, 2011). Botox is known to dampen muscle activity through neuromuscular paralysis and therefore also reduces facial muscle feedback and mimicry, whereas Restylane does not. The researchers found that the Botox group performed significantly poorer on the emotion recognition task compared to the Restylane group. Although there was a lack of random assignment in this study, emotional reactivity of participants was assessed prior to the study and found to be equal across both groups, leading to the interpretation that the observed emotion perception differences were most likely due to the inhibition of facial mimicry in the Botox group.

Further support for the role of mimicry in the understanding of emotion in others is provided by Sonnby–Borgström (2002), who found that participants displaying lower levels of

mimicry (as assessed by EMG) during a simple emotion identification task (happy/sad) also produced significantly lower scores on a self-report scale designed to assess emotional empathy. Although the emotion identification task wasn't designed to provide a measure of emotion perception ability (since the task was easy), it is implied that lower levels of empathy are a result of lower levels of mimicry impairing the full embodiment of the observed emotion.

      Facial mimicry and emotion perception in PD. The earlier mimicry work in neurotypicals by Niedenthal et al. (2001) and Oberman et al. (2007) that links production and perception deficits to one another via the mechanism of facial mimicry is consistent with recent findings from the present author and colleagues (Livingstone et al., 2016). Using EMG, we documented significantly lower levels of facial mimicry in a group of 27 PD patients when compared to age-matched healthy controls, with most pronounced deficits observed in mimicry responses to happy (zygomaticus) and sad (medial frontalis) facial expressions (compared to calm, angry, and fearful expressions). Furthermore, while the PD group's accuracy on the emotion perception task was comparable to the control group's, their response time when identifying happy emotions was delayed, and the extent of mimicry was negatively correlated with the extent of this delay.

      Another recent EMG study exploring facial mimicry and emotion perception in PD obtained similar findings. Argaud and colleagues (2016) presented 40 PD patients and 40 healthy controls with happy, angry and neutral facial expression avatars while recording muscle activity in various facial regions. The authors found that mimicry of happy expressions, as detected in the corresponding zygomaticus muscle, was significantly lower than in healthy controls. On an emotion perception task involving the same emotion presentations, PD participants identified

happy expressions with significantly less accuracy than healthy controls, but with no differences in the recognition of angry and neutral expressions.

It is interesting to note the somewhat conflicting findings of the above studies that revealed corresponding mimicry and emotion perception deficits specifically for happy expressions (Argaud et al., 2016; Oberman et al., 2007), and the findings of other PD studies that either noted least impairment in the perception of happy expressions (e.g., Assogna et al., 2010; Gray & Tickle-Degnen, 2010; Kan, Kawamura, Hasegawa, Mochizuki, & Nakamura, 2002; Sprengelmeyer et al., 2003; Suzuki et al., 2006) or failed to isolate any specific deficit in the perception of happy expressions (Livingstone et al., 2016). The most parsimonious explanation may lie in the earlier discussion of intensity of emotion, where it was noted that the majority of PD studies use happy stimuli that lacked the level of difficulty/intensity required to detect perceptual deficits in PD samples (Buxton et al., 2013). Indeed, the studies showing greater impairments in the perception of happy expressions have used either low-intensity static stimuli or dynamic presentations of expressions that require participants to identify the point at which an expression becomes distinct, whereas other studies have favoured stimuli with high accuracy ratings during validation (and therefore lower difficulty). The impairment in perception of happy expressions as a result of impaired mimicry may be somewhat counteracted by the distinct appearance of happy expressions (making this emotion easier to discriminate from other emotions through cognitive means), hence the need for lower intensity stimuli to detect any impairment.

Other evidence supporting the presence of deficits in spontaneous mimicry has been obtained using neural measures. Because spontaneous mimicry is thought to be the peripheral manifestation of activation of the mirror neuron system – a network of neurons in the brain that

activate both when an individual performs and observes a motor action – abnormal mirror neuron system activity would be expected to be present in PD patients. Indeed, Heida, Poppe, de Vos, van Putten, and van Vugt (2014) recently demonstrated such an abnormality in a pilot study of nine individuals with PD. Participants observed videos of hand movements while electroencephalography (EEG) was used to record event-related desynchronization of the mu wave, an EEG index of mirror neuron activity. Event-related desynchronization in response to videos was observed in the PD patients, suggesting that they had reduced mirror neuron system activity compared to healthy controls.

Vocal Characteristics of PD

As previously discussed, another important component of emotional communication is expression through the voice. Approximately 90% of PD patients experience vocal problems (Ho, Iansek, Marigliani, Bradshaw, & Gates, 1999), originating from the impairment of muscles involved in speech production, articulation and respiration (Skodda, 2012). Although dopamine replacement medications can improve vocal impairments, they nonetheless remain prominent (Pinto et al., 2004). Reduced volume, a harsh and breathy voice, shorter phonation duration, and limited intonation are often present during early stages of the disease and remain largely consistent throughout disease progression (Holmes, Oates, Phyland, & Hughes, 2000; Skodda, 2012). In later stages, fluency issues such as initiation difficulties, syllable repetition, pauses, and hurried spurts of speech also become apparent (Ho et al., 1999), with vocal tremors being particularly characteristic of advanced disease stage (Holmes et al., 2000).

Other common measures of vocal quality include jitter and shimmer, which are acoustic terms that refer to variations in fundamental frequency (pitch) and amplitude (loudness) (Teixeira, Oliveira, & Lopes, 2013). All vocalizations consist of cycles of sound waves, and it is

the properties of these cycles that determine how a voice sounds. Jitter specifically refers to the perturbations in fundamental frequency between sound wave cycles and is caused by irregularities in vocal fold vibrations (see Figure 1). Conversely, shimmer refers to perturbations in the amplitude between sound wave cycles and is associated with reduced resistance of the glottis (the opening between the vocal folds). Since jitter and shimmer are the result of interruptions to the vibrations of the vocal folds during exhalation, they are generally measured through tasks involving steady vocalization of a sustained vowel and with the use of acoustic software to extract a percentage value for each. Performance by PD patients on such tasks has revealed significantly higher levels of jitter and shimmer compared to healthy controls (Tanaka, Nishio, & Nimi, 2011). These higher levels of jitter and shimmer manifest perceptually as the cracked and hoarse voice quality that characterizes PD and, as mentioned earlier, do not appear to substantially worsen with disease progression.
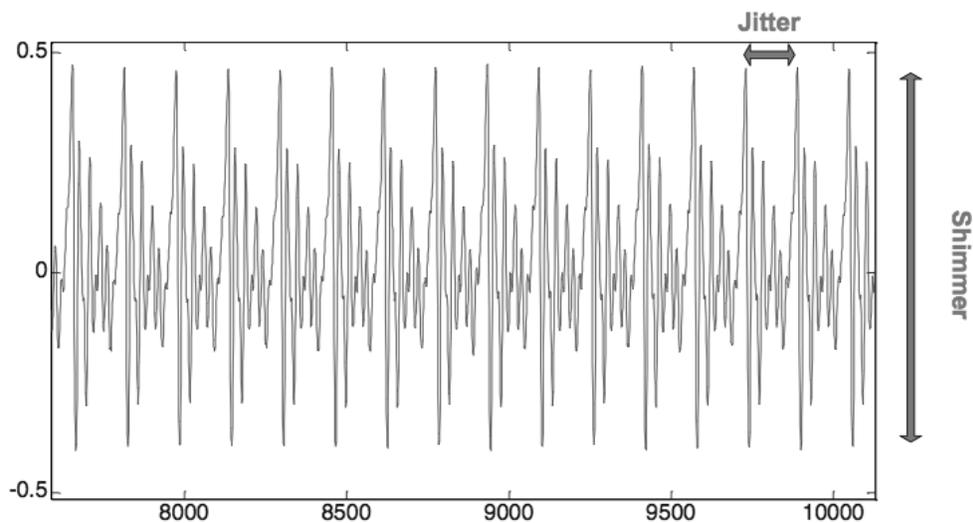
Figure 1. Graphic from Teixera et al. (2013) depicting jitter and shimmer as the result of perturbations in frequency and amplitude, respectively.

Interventions Targeting Vocal Deficits in PD

A number of treatments have been specifically designed to address vocal deficits in PD. Perhaps the most common and empirically supported of these is the Lee Silverman Voice Treatment (LSVT), a near-daily, four-week vocal exercise program which involves encouraging maximal effort during phonation and speech tasks and self-monitoring speech volume. The LSVT has been found to lead to improvements in volume and pitch variability, and was shown to be superior to an alternative treatment focused on promoting respiratory effort both at posttreatment and at two-year follow-up (Ramig et al., 2001). Overall speech intelligibility was also found to improve following LSVT (Cannito et al., 2012), with greater improvements even in comparison to articulation-focused treatments (Herd et al., 2012). However, it should be noted that a lack of comparative studies prevents the drawing of any firm conclusions about the superiority of any speech and language therapy over another (Herd et al., 2012).

Choral singing has also been shown to improve vocal characteristics of people with PD, though the literature is sparser and results appear to be mixed. For example, one exploratory study found improvements in pitch range following 20 weekly sessions of group singing, but no improvements in speech quality or most aspects of self-reported severity of vocal impairments (Elefant, Baker, Lotan, Lagesen, & Skeie, 2012). Another study investigating the effects of a two-year choir program consisting of fortnightly two-hour sessions also found improvements in pitch range, as well as speech volume and clarity of phonation (Evans, Canavan, Foy, Langford, & Proctor, 2012). At the lower end of treatment duration, a study by Stegemöller, Radig, Hibbing, Wingate, and Sapienza (2016) found that an eight-week group singing intervention consisting of a range of vocal exercises and singing of familiar songs led to significant improvements in respiratory pressure and phonation duration, but no significant changes in pitch

18

range or loudness. However, while a pilot study examining the effects of a 12-week choir also failed to find any changes in pitch range or loudness following the intervention, they additionally failed to find any improvements in phonation duration (Shih et al., 2012). Notwithstanding limitations such as sample size and use of a control group, these studies highlight the role of factors potentially responsible for differential results, e.g., length, intensity and content of the intervention.

As suggested above, most vocal interventions appear to fall into one of two categories: treatments administered one-on-one that involve a range of vocal exercises (such as the LSVT treatments), and treatments that are administered in a group format and combine vocal exercises with singing of songs. As well as requiring more time and resources to execute such individually-administered treatments on a regular basis, some researchers have pointed out the social benefits of a group singing format, which can lead to important improvements in mood and overall quality of life (Fogg-Rogers et al., 2016; Stegemöller et al., 2016), as well as encourage continued engagement and treatment compliance (Yinger & Lapointe, 2012). In addition, choirs, including those intended for people living with PD, are already quite popular in many communities, meaning that studies investigating choral singing trade tight control of variables for a more externally valid, applied and accessible approach.

Rhythmic Auditory Stimulation as an Intervention for Motor Deficits

Rhythmic auditory stimulation, i.e., the presence of an audible rhythm, has consistently been shown to aid intentional gross motor movements in PD patients, leading to demonstrable improvements in gait (e.g., de Bruin et al., 2010; Hackney, Kantorovich, Levin, & Earhart, 2007; Ledger, Galvin, Lynch, & Stokes, 2008; Thaut et al., 1996). The theory of rhythmic auditory stimulation in this context stipulates that deliberately walking in time to an audible beat aids the

gross motor action of walking by accessing motor areas of the brain via neural networks that bypass the areas affected in PD. The rhythm is believed to compensate for the time-keeping dysfunction of the basal ganglia, which is present due to the destruction of dopamine-producing cells as a result of the disease (McIntosh, Thaut, Rice, & Miller, 1994).

In terms of music-based interventions involving rhythmic auditory stimulation, Pacchetti et al. (2000) found that a three-month course of weekly music therapy involving choral singing, vocal exercises and free movement led to improvements in several gross motor functions. Given that the brain entrains to rhythm (Thaut, Kenyon, Schauer, & McIntosh, 1999), it is likely that the implicit rhythm within song also facilitates execution of facial and vocal movements during singing, which may eventually retrain damaged neural circuitry involved in the expression of emotion. In one of the few studies to focus on rehabilitating observable facial expressiveness in PD, a Norwegian research group (Elefant et al., 2012) used the Facial Action Coding System (FACS; Ekman & Friesen, 1976) to rate the expressiveness of PD patients based on 15-second video recordings taken across 20 weekly music therapy sessions that included singing, vocal exercises and movement exercises. Although a relatively small sample was used ($N = 10$), the results revealed significantly improved natural facial expressiveness while singing and speaking during therapy sessions from the beginning to the end of treatment.

A recently completed study in our lab that is the precursor to the present study investigated a novel 13-week individually-administered protocol that combines rhythmic auditory stimulation with deliberate imitation of observed facial expressions (Livingstone, Vezer, McGarry, Lang, & Russo, in progress). In this study, a treatment PD group intentionally mimicked video clips of actors singing with various emotions (calm, happy, sad, fearful, and angry) while a control PD group passively observed the same videos. Following completion of

the protocol, the treatment group significantly improved facial mimicry responses to calm, happy, sad, and fearful expressions, as measured by EMG in the zygomaticus and medial frontalis muscle regions. Since surprised and disgusted expressions were not targeted in the intervention, the findings suggest that the effects of such expressive exercises on facial mimicry do not generalize to the mimicry of emotions that are not specifically targeted.

It is worth noting that the video clips alternated between two seven-syllable sentences which, when sung, were set to a very basic melody (Livingstone et al., in progress), making the exercise highly repetitive and not particularly engaging. Although this format was necessary in the initial stages of experimentation, promising results of the intervention support its modification to a more patient-friendly format and also suggest that gains in facial mimicry during singing interventions may arise independent of a strong social component and task engagement. An alternative to this delivery format is choral singing, which, as discussed in a previous section, not only offers additional social benefits and higher levels of engagement, but is also more cost-effective and more reflective of the kinds of programs that are readily available to the public.

Effects of Vocal Training on Facial Expressiveness

Even in the absence of rhythmic auditory stimulation in vocal treatments, the underlying neural mechanism that is shared by both facial and vocal muscle activity (Jürgens, 2002) and the pairing of the two during vocal activities mean that treatments aimed at addressing vocal deficits in PD may also inadvertently address impairments in facial expressiveness. Spielman et al. (2003) investigated this theory by measuring the effects of the LSVT LOUD (Ramig, Fox, & Sapir, 2008) – an intensive treatment that focuses on making voices louder – on the natural, involuntary facial expressiveness of a group of PD patients, half of whom underwent the

21

treatment while the other half did not. Prior to and following each treatment condition, trained

independent raters who were blind to the assigned conditions evaluated participants' facial

mobility while they talked about a chosen topic for 25-30 seconds. The authors found a

significantly greater increase in facial mobility following the treatment condition, supporting the

hypothesis that vocal exercises alone can also positively impact facial muscle activity.

A more recent study by Dumer et al. (2014) attempted to elucidate the relationship

between vocal treatments and improvements in facial expressiveness by extending Spielman and

colleagues' study with the addition of a PD group that underwent vocal articulation treatment

instead of LSVT LOUD and a group of healthy controls. In this study, participants in the LSVT

LOUD group, the articulation treatment group, the non-treatment group, and the healthy control

group all produced one-minute monologues about a happy personal experience, both before and

after the treatment period. The use of recalling and talking about happy personal experiences is

based on evidence that happiness tends to evoke more facial expressiveness (Smith et al., 1996).

On measures of facial activity frequency and variability, independent ratings revealed clear

posttreatment gains in the LSVT LOUD group, but not the articulation therapy group.

Furthermore, despite all PD groups demonstrating lower facial activity levels at pretreatment

than healthy controls, the LSVT LOUD group improved enough on these measures at

posttreatment to resemble facial activity levels of the healthy control group.

Dumer and colleagues (2014) speculated that neuroimaging studies may shed light on the

potential mechanism responsible for these findings. Studies by Liotti et al. (2003) and Narayana

et al. (2010) suggest that LSVT training may strengthen parts of the brain that are involved in

controlling amplitude of motor action. Accordingly, it is possible that any activity that focuses on

amplitude (as in the case of LSVT LOUD, which is designed to increase the vibrational

amplitude of the vocal folds) also increases the amplitude of facial muscle movements by strengthening the neural structures that are responsible for both, whereas a focus on clear speech (as in the case of articulation therapy) does not.

The Present Study

As an extension of our previous study, the present study sought to investigate the impact of a choir program on specific aspects of emotional communication. The choir program was specifically developed to target happy and sad emotional expressions, since these two emotions were shown to be most impaired in terms of facial mimicry in our previous study (Livingstone et al., 2016) and because happiness and sadness – unlike other emotions – are commonly and easily conveyed through song. Given the link between facial mimicry and emotion perception, perception of happy and sad emotions would therefore also be targeted. In addition, choral singing directly engages the vocal muscles and may consequently address vocal expressiveness deficits in people with PD.

The choir program was intended for anyone with PD wishing to take part in it (with the additional option of taking part in the present research) and was organized in collaboration with the Royal Conservatory of Music in Toronto, Ontario. Recruitment of members was organized through Parkinson Society Canada (PSC) support groups, with whom we have had previous partnerships. Flyers with brief information about the choir and the related research opportunity had been distributed by PSC on our behalf to the 15 support groups around the Greater Toronto Area via e-mail and also placed in paper form on the head office's notice board. We also gave a number of brief in-person presentations to support groups.

Prior to and immediately following participation in a 13-week choir program, EMG was used to measure facial mimicry during observation of videos of expressed emotion. Specifically,

activity in the zygomaticus and medial frontalis muscles was measured, since these were the

muscles in which mimicry impairments were found in our previous study (Livingstone et al., in

progress). A separate emotion perception task sought to determine participants' ability to identify

emotion in facial expressions. Although sad and happy emotions were targeted for improvement,

the mimicry and emotion perception tasks also included calm, angry and fearful emotions in

order to rule out any generalized effects beyond those of the targeted emotions. In addition, vocal

testing detected a range of vocal features that are most consistently impaired throughout all

stages of disease progression in PD (lowest and highest achievable pitch, loudness, phonation

duration, jitter, and shimmer).

Method

Participants

Prior to recruitment, *G*Power* was used to conduct power analyses to determine the target sample size. Aside from our previous study (Livingstone et al., in progress), there were no studies with comparable research questions and measures that also reported effect sizes, therefore it was not possible to conduct power analyses for each measure. As a result, power analyses were conducted based on an effect size of .48 yielded in our previous study for our most important statistical test: a repeated measures t-test comparing pre and postchoir timepoints for zygomaticus muscle activity in response to happy expressions. Based on $\alpha = .05$ and $\beta = .80$, the power analysis determined that 37 participants would be needed.

During the first choir session, group members learned about the research component of the program via a brief talk and were informed that the purpose of the study was to investigate the effects of choir participation on emotional communication in PD. Information provided about the study included the procedure, potential risks and benefits and eligibility criteria. After being given the opportunity to answer questions, choir members received contact details of the researchers and asked to make contact after the choir session if they were interested in participating or if they had any further questions. It was made clear that participation in the study would not affect individuals' choir membership. Of the 21 choir members, only 12 individuals agreed to take part in the research. Since this number fell short of the required sample size of 37, the study was likely underpowered.

Eligibility

Participants were required to be between the ages of 50-75, have been diagnosed with idiopathic PD (i.e., the cause of the condition is unknown) over the age of 50, have no other

diseases of locomotion, not have recently participated in or currently be participating in singing-based programs, and be at a Hoehn and Yahr stage of between 2 and 5. The Hoehn and Yahr stages are part of a rating scale commonly used as a measure of disease progression in PD, in which stages 2-5 can be considered moderate-to-severe levels of progression (Goetz et al., 2004). The lower end of this range of stages is characterized by disease affecting both sides of the body and some postural instability, but overall intact physical independence, while the higher end is associated with an inability to stand or walk without assistance. Although this range of severity is arguably wider than would be ideal, we felt it necessary at this early stage of research to adopt more inclusive eligibility criteria, and also in consideration of the study timeframes and recruitment challenges common in this population. The above eligibility information was self-reported by participants and, in the case of Hoehn and Yahr stage, was easily obtained by participants from their neurologists.

All participants were actively taking dopamine replacement medication for PD (levodopa), and were advised to schedule testing for a time when the full effects of the medication could be expected to endure for the entirety of the testing period (2.5-3 hours), or as close as possible. As is generally the case, medication types and dosages varied between patients; however, all participants reported remaining on a consistent dosage of medication throughout the study period.

Since depression has been found to have independent effects on emotion perception (see Dalili, Penton-Voak, Harmer, & Munafò, 2015) participants were also required to complete the Beck Depression Inventory-II (BDI-II; Beck, Steer, Ball, & Ranieri, 1996), a 21-item self-report rating scale designed to measure symptoms and other characteristics of depression. Participants scoring in the range of severe depression (defined as a score of 28 or more) or endorsing items

related to suicidal ideation were to be excluded from the study and provided with appropriate

resources for dealing with depression/suicidal thoughts. However, no potential participants scored

beyond the mild range (defined as scores from 14-19). The Montreal Cognitive Assessment (MoCA;

Nasreddine et al., 2005) was also administered in order to rule out dementia (defined as a score of 21

or less) in order to ensure adequate cognitive ability to complete required tasks. All participants were

also required to have self-reported normal or corrected-to-normal hearing and vision. See Table 1

below for participant characteristics.

Table 1

Participant Characteristics

Variable	N
Male	8
Female	4
Medicated	12
Variable	M±SD
Age (years)	69.58±9.00
Years since diagnosis	4.75±2.38
MoCA	27.58±1.31
BDI-II	8.75±5.17
Choir sessions attended	10±1.21
Variable	Mdn (range)
Hoehn and Yahr stage	2 (2-5)

Note. M = Mean, SE = Standard Error, Mdn = Median, MoCA = Montreal
Cognitive Assessment, BDI-II = Beck Depression Inventory-II

Choir Program Curriculum

The 13-week choir curriculum was informed by the findings of our previous study

(Livingstone et al., in progress), and was led by a professional choir director with

accompaniment provided by a trained pianist. Each weekly session consisted of 10 minutes of

vocal exercises, such as breathing exercises, pitch range exercises, and note sustenance exercises,

followed by 40 minutes of learning and practicing selected songs. In total, the choir developed a

repertoire of 10 songs that were introduced gradually throughout the 13 weeks. A novel and critical aspect of the choir program is that the choir director modeled happy and sad expressions during singing, with choir members instructed to imitate the smiles and frowns that were consistent with the emotions being conveyed in the songs. As previously discussed, the use of specifically happy and sad songs was based on our findings that these are most impaired in terms of facial mimicry, and are also the two basic emotions that are most commonly and easily conveyed through music. Selected songs were also expected to encourage genuine emotional expression by eliciting felt emotion. The songs were selected in collaboration with choir members, and included songs such as Empty Chairs at Empty Tables from the musical *Les Miserables* and Good Morning Starshine from *Hair*.

Stimuli

The RAVDESS (Livingstone, Peck, & Russo, 2012) is an extensive database of over 7000 high quality audio, video and audio-video clips featuring 24 professional actors (Caucasian, 12 male/12 female) singing and speaking statements with a range of emotions (see Figure 2). The statements "Dogs are sitting by the door" and "Kids are talking by the door" are sung in a basic melody or spoken, and were selected for their prosodic properties and neutral meaning. Emotions displayed by the actors are calm, happy, sad, angry, and fearful. Each emotion is conveyed in the above sentences through speech, song with metronome, and song without metronome, ranging in duration from 4-6 seconds. The database includes accuracy averages gathered during the authors' validation process, meaning that videos can be selected from the database based on difficulty level of emotion identification. These dynamic, multimodal stimuli were selected for their closer approximation of real life examples of facial expressiveness in interpersonal contexts compared to static images.

Figure 2. Samples of RAVDESS actors singing and speaking with emotional expressions

Apparatus and Materials

Electromyography. EMG was used to record facial muscle activity for the purposes of measuring facial mimicry before and after choir participation. Specifically, a BioPac MP150 acquisition system along with a BioNomadix 2Ch Wireless EMG amplifier was used to collect data. As older adults and patients with PD can suffer from dry skin conditions, a detailed three-step skin preparation process was followed. NuPrep gel was first applied to electrode areas in order to lower contact impedance in the zygomaticus and medial frontalis muscle regions, and symmetrically on the mastoid bones behind each ear. After five minutes, outer layer non-conductive skin cells were removed with the use of 20x25mm abrasive pads. Finally, a small amount of NuPrep gel was reapplied to all regions to smooth the skin. Pairs of 25mm square cloth Ag/AgCl electrodes (11 mm contact area) were then applied to zygomaticus (smiling) and medial frontalis (frowning) muscle regions on the left side of the face, while two further electrodes were placed on each mastoid bone behind the ears to act as a voltage referent. Placement of electrodes on the left side of the face is common practice in facial mimicry studies due to preferential processing of emotional stimuli in the right hemisphere of the brain resulting in somewhat stronger expressions on the left side of the face (Dimberg & Petterson, 2000; Zhou & Hu, 2004). Electrodes were wired to the BioNomadix transmitter, which communicates wirelessly with the BioPac MP150W system. EMG signals were digitized with AcqKnowledge

4.2 software. Participants were tested individually in a testing booth at the Institute for Stress and Wellbeing Research at Ryerson University, where stimuli for the facial mimicry task were presented on a 22-inch LCD display monitor and Logitech speakers, placed 60cm from the participant. Stimulus presentation was controlled with custom software written in Presentation while EMG data were collected throughout and saved onto a Fujitsu Lifebook computer.

Emotion perception task. For this task, RAVDESS stimuli were presented via QuickTime software on an 11-inch mid-2012 MacBook Air running OS X version 10.9.5. Emotion options were presented on a printed sheet of letter-sized paper for participants' reference and verbal responses noted by the experimenter on a separate sheet of paper out of view of the participant.

Vocal recordings. Participants were audio recorded in a sound-attenuated recording booth while performing a range of vocal tasks that are described in the next section. Recording was carried out in ProTools for iMac and with the use of a height-adjustable Rode NTK microphone, pop-filter and tripod, and with input settings consistent across all trials. A standard tape measure was used to position the microphone 88cm from each participant's mouth.

Procedure

Choir members who made contact with the research team to express interest in participation were sent a copy of the consent form via e-mail or mail for review and asked to obtain their most recently assessed Hoehn and Yahr stage from their neurologist for eligibility purposes. At this point, participants were also asked questions to establish eligibility criteria regarding age, vision and hearing, medication regimen, language ability and participation in other singing-based programs. The first testing session was scheduled once these preliminary eligibility criteria were met and were done so for a time when each participant expected to be in

their post-medication "on" state for as much of the testing duration as possible. All participants were able to be scheduled for testing within the first two weeks of choir participation. During the first testing session, informed consent was obtained and the MoCA and BDI-II administered to check for remaining eligibility criteria. All individuals met full eligibility criteria and, as mentioned earlier, none were identified as having severe depression or endorsing suicidal ideation. Participants then progressed to the testing portion of this session as per the following procedures, which were repeated within two weeks of completion of the choir program. An honorarium of $30 was provided to each participant upon completion of both testing sessions. All 12 participants completed both components of testing.

Facial mimicry task. Participants watched a series of videos of actors singing and speaking with a range of emotional expressions that were selected from the RAVDESS (Livingstone et al., 2012). Each participant observed a total of 120 videos, consisting of all combinations of emotion (calm/happy/sad/angry/fearful), sex of actor (male/female), sentence ("Dogs are sitting by the door"/"Kids are talking by the door"), modality of presentation (song with metronome/song without metronome/speech)[1], and two repetitions. Stimuli were presented in blocks based on modality and counterbalanced across participants, with all other factors presented randomly.

During observation of the videos, participants were asked to specify the emotion observed in each video using a computer keyboard with number keys corresponding to the five emotion options. The presentation of each clip was preceded by a 100 ms beep followed by 100 ms of silence, during which time a black screen with a centred fixation point was

---

[1] Although the effects of modality are not relevant to the research questions we wish to address – and as such we have no specific predictions in relation to it – it was nonetheless included in the design to test for potential differential effects given that the intervention involved song.

shown. The stimulus was then shown (duration of 400-600 ms), after which a black screen with the five possible key presses appeared along with the instructions "Please choose the correct emotion." The next video clip was not played until the participant made a selection. The duration of the task was approximately 45 minutes. Throughout this task, EMG was used to record facial muscle activity of participants, providing spontaneous facial mimicry data. The emotion identification task in this trial was intended only to keep participants engaged with the testing process, and videos with higher accuracy values during validation (80% and above) were selected for this purpose in order to elicit stronger and more detectable mimicry responses.

      Emotion perception task. Accuracy of identification of emotional facial expressions was measured in a separate task by presenting video clips taken from the RAVDESS of actors expressing the five emotions of interest through speech only, and asking participants to identify the target emotion. A total of 30 videos consisting of six presentations of calm, happy, sad, angry, and fearful emotions performed by an equal number of male and female actors were shown, with participants telling the experimenter which emotion they believed was being conveyed before the next video was played. A letter-sized printout of the five emotion options was placed in front of participants for reference, and responses were noted by the experimenter. Importantly, videos that had achieved accuracy averages of 50-70%[2] during validation of the stimulus database were specifically selected for this task in order to avoid a ceiling effect, and each participant was presented with one of six random sequences of the same videos at each testing timepoint.

---

[2] An accuracy average of 50% does not represent equal chance, since the validation process included seven emotion options.

Vocal tasks. For all vocal tasks, the microphone was placed in a uniform position at one end of the sound-attenuated recording room, with the microphone itself adjusted in height on the tripod to be in line with each participant's mouth. Participants were asked to stand so that their mouths were 88 inches from the microphone, guided by a marking on the floor and verified by measuring the distance with a tape measure. Once these adjustments were made, recording began and the following tasks were carried out.

Pitch range task. Participants were asked to vocalize the vowel "Ah" from the lowest possible note they were able to vocalize to the highest possible note, either in the form of a glissando (continuous upward slide of notes) or by repeating "Ah" in an ascending scale. Participants were asked to attempt this task a second time with greater effort. This task provided measures of lowest and highest achievable pitch.

Loudness task. Participants were asked to vocalize the vowel "Ah" as loudly as they could, followed by a second trial in which they were asked to attempt to be even louder. This task provided the measure of maximum loudness.

Phonation duration task. Participants were asked to take a deep breath and sustain the vowel "Ah" for as long as possible, followed by a second attempt. This task was used to measure the maximum duration of the sustained vowel, or phonation, as well as jitter and shimmer.

Data Analysis

Data were analyzed for each measure as described below. Additionally, analyses were run to detect any between-subjects effects of sex or any covariances with MoCA scores, BDI-II scores, Hoehn & Yahr stages, and choir attendance. In all cases, no significant effects or relationships were found ($p > .05$) and all effect sizes (in Pearson's r) were below the threshold for a small effect size (Cohen, 1988; Cohen 1992).

Facial mimicry data. EMG data from the facial mimicry task were filtered using a notch filter to attenuate ambient electrical noise (Butterworth, second order, 59–61 Hz), and then high-pass filtered to minimize motion artifacts (Butterworth, third order, 30 Hz cut-off) and low-pass filtered to minimize high-frequency noise artifacts (Butterworth, third order, 400 Hz cut-off). Data were full-wave rectified and then smoothed using an RMS filter with a 50 ms sliding window and overlap of 49 ms. Data were also zeroed using a baseline subtraction procedure. A baseline window of 2000 ms prior to video onset was selected. The average value within the baseline window was subtracted from the target window. The target window is defined as 3000 ms, which is the duration spanning the start and end of the video presentations of the actors.

Separate repeated measures analyses of variance (ANOVAs) of mean amplitude of muscle activity (in μV) were carried out for the zygomaticus and medial frontalis muscles in response to all five emotions. While all emotions were included in each analysis, the emotions of interest for each muscle are those that have been shown to elicit activity in that particular muscle (i.e., calm and happy for zygomaticus, and sad, angry and fearful for medial frontalis). Factors and levels included timepoint (pre/post) X emotion (calm/happy/sad/angry/fearful) X modality (song with metronome/song without metronome/speech) X repetitions (8), with only timepoint and emotion being of interest to study hypotheses. Further, separate repeated measures ANOVAs were conducted only for happy and sad emotions (based on zygomaticus and medial frontalis muscle activity, respectively) to test our differential predictions about the pre and postchoir effects for these two specifically targeted emotions.

Emotion perception data. Data from the emotion perception task were generated in binary format (correct/incorrect), therefore a value of 0 or 1 was assigned to incorrect and correct answers respectively, providing an accuracy score between 0 and 6 for each emotion (as the task

consisted of six displays of each emotion). Data were then analyzed using a repeated measures ANOVA with the factors timepoint (pre/post) X emotion (calm/happy/sad/angry/fearful). Given our specific hypotheses about the effects of the choir on the perception of happy and sad expressions specifically, a paired-sample t-test was run for each of these emotions.

Vocal data. Since each vocal task was performed twice, the best of the two attempts within each task was selected for analysis (e.g., the pitch range attempt with the highest/lowest pitch, the phonation duration attempt with the longest duration, etc.). Acoustic features were extracted from vocal recordings with the use of version 5.3.55 of the acoustic analysis software PRAAT using the methods described below. Frequency and variability distributions for much of vocal data violated assumptions of normality and homogeneity of variance, therefore nonparametric robust tests were used for analysis. Specifically, separate Wilcoxon signed-rank tests were carried out for each vocal feature, with medians reported rather than means and standard errors.

Pitch range task data. Since participants were asked to vocalize from the lowest possible pitch to the highest possible pitch, lowest achieved pitch was calculated (in Hz) within the first three seconds from vocalization onset while highest achievable pitch was calculated using the final three seconds prior to termination of the vocalization. This process is consistent with other studies using this measure (e.g., Holmes et al., 2000).

Loudness task data. Maximum amplitude values (in dB) were calculated for the entire vocalization, regardless of length.

Phonation duration task data. The period of time between vocalization onset and termination was calculated to determine vowel sustain duration (in seconds). Following Holmes et al. (2000) and Tanaka et al. (2011), jitter and shimmer were calculated based on a three-

second window in the middle portion of the vocalizations produced in the phonation duration task.

Hypotheses

Hypotheses about the effects of the choir program on the various outcome measures are as follows.

Hypothesis 1: Mimicry task. Participants will demonstrate significantly higher levels of activity in the zygomaticus muscle region in response to observations of happy emotions, but not calm emotions, since choir songs targeted happy expressions only and our recent study suggests that only targeted emotions improve corresponding mimicry (Livingstone et al., in progress). Similarly, higher levels of activity in the medial frontalis muscle region will be observed following choir participation in response to sad emotions, but not angry or fearful emotions, as only sad emotions were targeted by the choir's repertoire of songs.

Hypothesis 2: Emotion perception task. Participants will improve accuracy scores for identification of happy and sad emotions following choir participation, but not calm, angry, or fearful, since the former were the only two emotions to be targeted by the choir songs.

Hypothesis 3: Vocal tasks. Given the inherent emphasis on breath control, loudness and pitch modulation in choral singing, as well as the general strengthening of the vocal apparatus, we expected that choir participation will result in: (1) an increase in phonation duration; (2) a decrease in lowest achievable pitch; (3) an increase in highest achievable pitch; (4) an increase in maximum loudness; (5) a reduction in jitter levels, and; (6) a reduction in shimmer levels.

Results

Because of limitations in power, analyses for mimicry data were restricted to omnibus tests for each muscle group followed by targeted tests to assess specific emotion-muscle training hypotheses. Effect sizes are expressed in Pearson's *r* for all comparisons involving prechoir/postchoir differences, whereas η_p^2 is used for any comparison involving multiple groups and interactions.

Preliminary analyses were conducted to check for the expected patterns of muscle activation (i.e., mimicry) in response to specific emotional expressions. For the zygomaticus muscle, these analyses confirmed that, as would be expected for this muscle, a main effect of emotion was found, $F(4, 40) = 7.904, p < .001, \eta_p^2 = .44$, indicating variability in activation of this muscle in response to presentations of the five different emotions. There was no overall effect of timepoint, $F(1, 10) = 2.727, p = .130, r = .46$, nor a significant interaction between timepoint and emotion, $F(4, 40) = 0.637, p = .639, \eta_p^2 = .06$. No main effects or interactions involving modality were found. For the frontalis muscle, a main effect of emotion was found, $F(4, 40) = 7.680, p < .001, \eta_p^2 = .43$, demonstrating differences in facial muscle activation in response to presentations of the five different emotions, as expected. There was no overall effect of timepoint, $F(1, 10) = 0.006, p = .939, r = .02$, but a significant interaction between timepoint and emotion was found, $F(4, 40) = 3.358, p = .018, \eta_p^2 = .25$. Visual inspection of plots showed that while muscle activity increased in response to sad and angry emotions at postchoir, it decreased in response to calm, happy, and fearful expressions. No main effects or interactions involving modality were found.

Mimicry Results

Zygomaticus muscle. Figure 3a plots zygomaticus muscle activation in response to calm and happy emotions as a function of emotion and timepoint. As predicted, an ANOVA only for happy emotions (timepoint X modality X repetitions) showed that zygomaticus activity in response to observation of happy expressions was marginally significantly greater following choir participation ($M = 0.268$, $SE = 0.095$) than prior to choir participation ($M = 0.127$, $SE = 0.064$), $F(1, 10) = 4.975$, $p = .050$, $r = .58$.

Medial frontalis muscle. Figure 3b plots medial frontalis muscle activation for sad, angry and fearful emotions as a function of emotion and timepoint. Contrary to expectation, an ANOVA only for sad emotions (timepoint X modality X repetitions) revealed that participants did not demonstrate significant changes in medial frontalis muscle activity in response to this emotion between prechoir ($M = 0.040$, $SE = 0.115$) and postchoir ($M = 0.120$, $SE = 0.123$), $F(1, 10) = 2.007$, $p = .187$, $r = .41$.
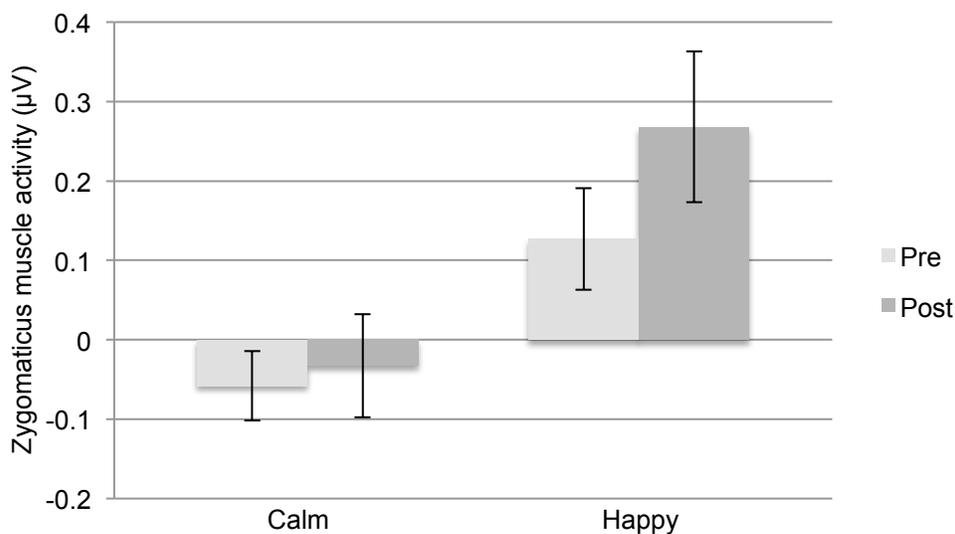
Figure 3a: Participants' mean zygomaticus muscle activity in response to displays of calm and happy expressions prior to (Pre) and following (Post) choir participation, with standard error bars.
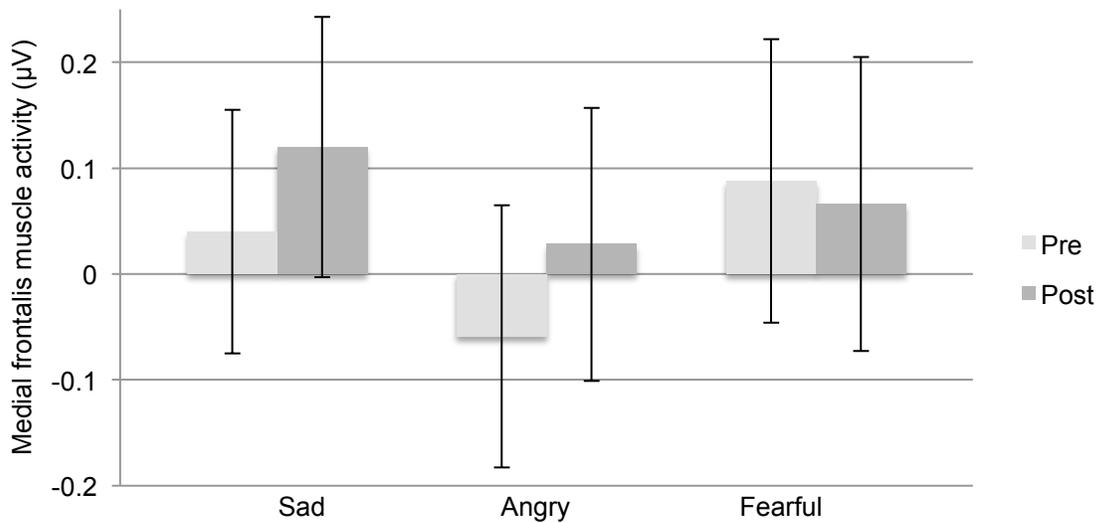
Figure 3b: Participants' mean medial frontalis muscle activity in response to displays of sad, angry and fearful expressions prior to (Pre) and following (Post) choir participation, with standard error bars.

Emotion Perception Results

Figure 4 plots accuracy scores on the emotion perception task as a function of emotion and timepoint. A significant main effect of emotion was found, $F(4, 44) = 12.562, p < .001, \eta_p^2 = .53$, suggesting that some emotions were easier to decode than others. There was no main effect of timepoint, $F(1, 11) = 1.084, p = .320, r = .31$, nor an interaction between timepoint and emotion, $F(4, 44) = 0.844, p = .505, \eta_p^2 = .08$. Accuracy scores for the perception of happy expressions before choir participation ($M = 4.333, SD = 0.482$) increased marginally significantly after choir participation ($M = 5.000, SD = 0.348$), $t(11) = -1.876, p = .087, r = .49$. However, accuracy scores for sad expressions at prechoir ($M = 2.417, SD = 0.358$) were not

significantly different compared to postchoir ($M = 2.167, SD = 0.441$), $t(11) = 0.761, p = .463, r$

$= .22$.
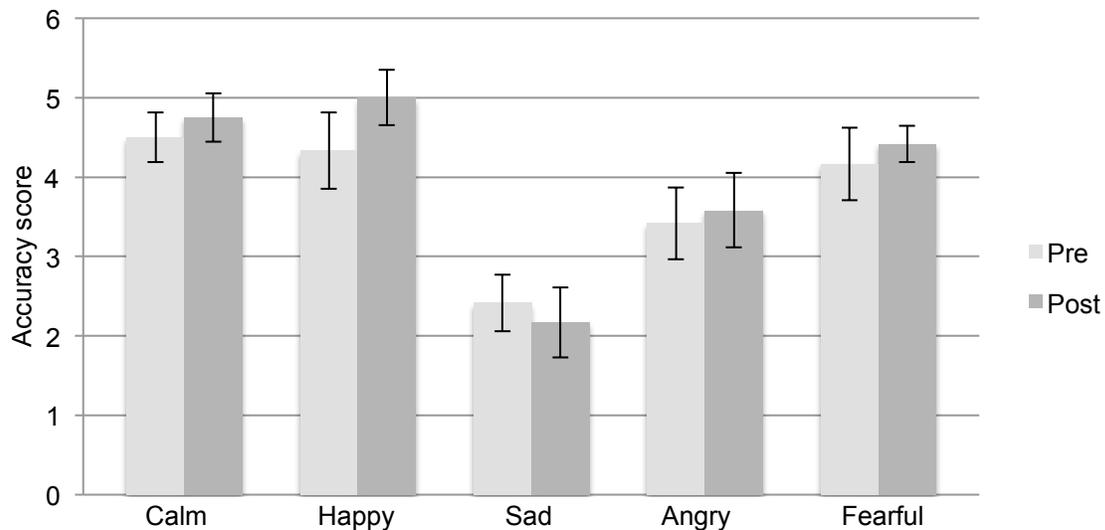
Figure 4: Participants' mean accuracy scores (out of six trials) on the emotion perception task before (Pre) and after (Post) choir participation, with standard error bars.

Vocal Analysis Results

Pitch range task. Figure 5a and 5b show participants' lowest and highest achievable pitch, respectively, before and after choir participation. Wilcoxon signed-rank tests showed that participants achieved a significantly lower pitch on the pitch range task following choir participation ($Mdn = 101.605$ Hz) compared to prior to participation ($M = 122.156$), $T = 3, p =$ $.005, r = .58$. However, there was no significant difference between the highest achievable pitch before choir participation ($Mdn = 431.331$) and afterwards ($Mdn = 410.1630$), $T = 51, p = .347, r$ $= .19$.
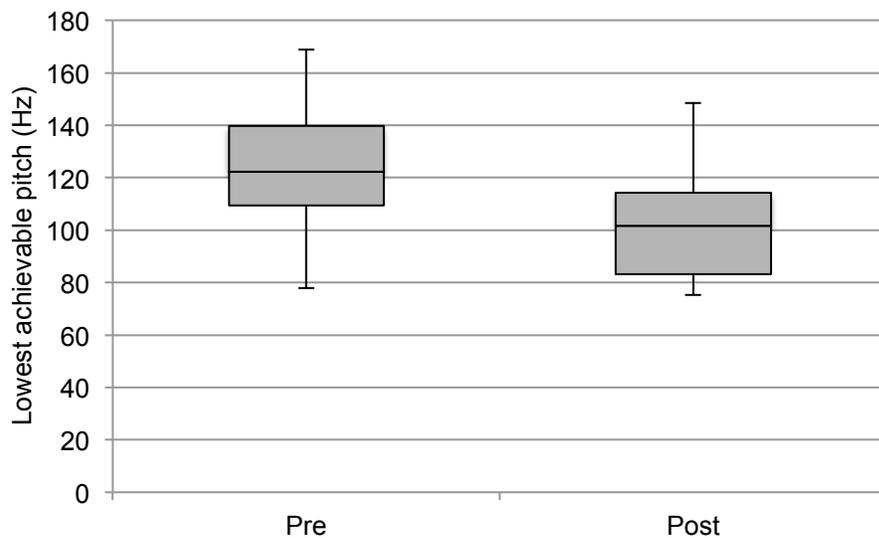
Figure 5a: Median lowest pitch achieved by participants before (Pre) and after (Post) choir

participation. Box = 25th and 75th percentiles, whiskers = minimum and maximum values.
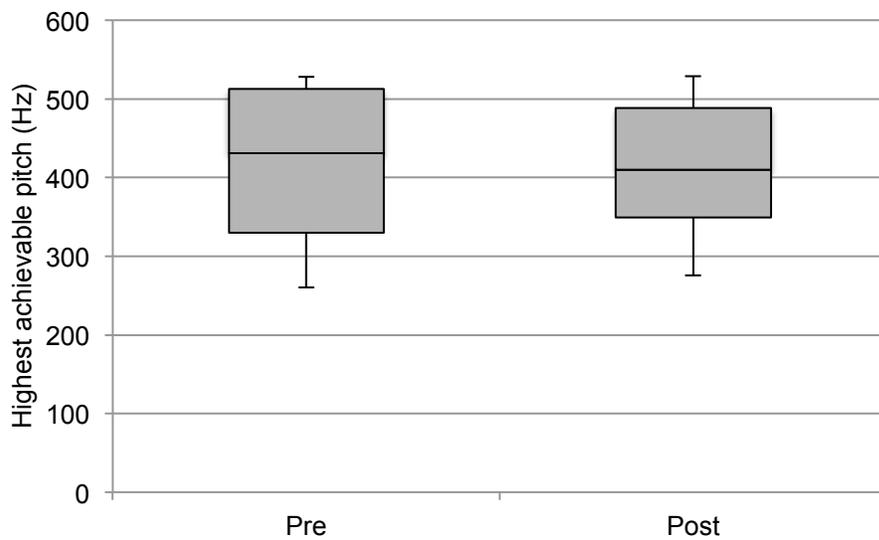
Figure 5b: Median highest pitch achieved by participants before (Pre) and after (Post) choir

participation. Box = 25th and 75th percentiles, whiskers = minimum and maximum values.

Loudness task. Participants failed to display any increase in maximum amplitude (in dB) following choir participation (*Mdn* = 103.991) compared to prior to participation (*Mdn* = 103.531), *T* = 37, *p* = .875, *r* = .03.

Phonation duration task. Although the duration (in s) for which participants could sustain the vowel sound "Ah" was longer following choir participation (*Mdn* = 18.182) than prior to participation (*Mdn* = 14.882), the difference only reached marginal levels of significance, *T* = 63, *p* = .060, *r* = .38. Jitter percentages during this task significantly decreased from prechoir (*Mdn* = 0.58) to postchoir (*Mdn* = 0.25), *T* = 13, *p* = .041, *r* = .42 (see Figure 6), while shimmer levels decreased (*Mdn* = 16.96 at prechoir, *Mdn* = 11.39 at postchoir) but only reached marginal levels of significance, *T* = 16, *p* = .071, *r* = .37.
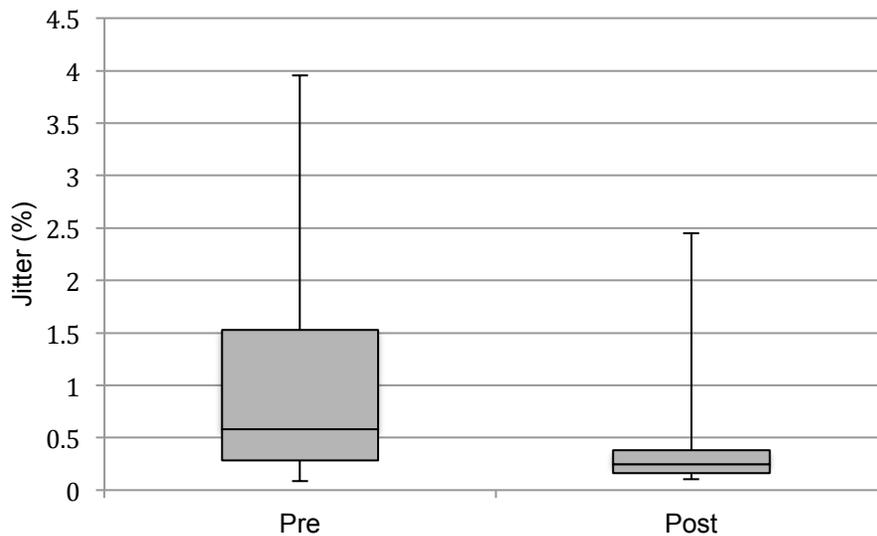

Figure 6: Median jitter percentages before (Pre) and after (Post) choir participation. Box = 25th and 75th percentiles, whiskers = minimum and maximum values.

Discussion

Facial Mimicry

This study was designed to investigate the rehabilitative potential of a specialized 13-week choir program for a range of emotional communication deficits in people with PD, specifically in the areas of facial mimicry, emotion perception and vocal quality. Based on the benefits of our previous study's singing therapy on the mimicry of specifically targeted emotions (Livingstone et al., in progress), and the targeting of happy and sad emotions in the choir program used for the present study, we predicted that mimicry of happy and sad expressions would significantly improve following completion of the choir, and that improvements in vocal parameters would also be observed. Despite the direction of change between prechoir and postchoir being in the hypothesized direction for the mimicry of both of the emotions targeted by the choir, neither mimicry of happy nor sad expressions reached statistical significance, although happy mimicry marginally significantly improved. Our initial power analyses suggest that a larger sample may have yielded more significant results, particularly in light of the finding for happy mimicry yielding a large effect size ($r = .58$) and the sad mimicry finding yielding a medium-to-large effect size ($r = .41$) (Cohen, 1988; Cohen, 1992).

Other plausible explanations for the lower magnitude of effect on the mimicry of sad expressions may lie in the content of the choir program. In developing the choir, we had proposed that the advantage of singing happy and sad songs is that music can induce emotion and therefore facilitate the production of genuine facial expressions congruent with the corresponding emotion. However, for many individuals, singing is in itself an enjoyable activity that induces positive emotion. This could mean that while the enjoyable act of singing likely facilitated felt emotion (and subsequent expression) when singing happy songs, it may have

overridden any sad mood induced by the singing of sad songs and therefore suppressed any corresponding sad facial expressions. That is, if participants enjoyed singing sad songs, they may have inadvertently displayed happy expressions rather than sad expressions, or displayed sad expressions that were forced instead of genuine. This would naturally result in a greater strengthening of the muscles required for happy expressions (and their mimicry) rather than sad expressions (and their mimicry). In addition, it is possible that the choir director encouraged happy expressions more than sad ones due to the aversive nature of sadness, and, although the number of happy and sad songs were equal, may have also spent more time on happy songs for similar reasons. It would have been useful in this regard to track the choir director's fidelity to researchers' instructions, as well as the proportion of time spent singing happy versus sad songs.

It is also possible that felt emotion is beneficial to the type of facial expression production that strengthens mimicry. There is evidence that genuine, felt emotion results in facial expressions that are visually distinct from voluntary or non-genuine expressions (e.g., Ekman, Davidson, & Friesen, 1990), which is consistent with the presence of discrete underlying neural circuitry discussed in a previous section (Rinn, 1984). Since involuntary facial expressiveness appears to be considerably more impaired in PD than voluntary expressiveness (Simons et al., 2003; Simons et al., 2004; Smith et al.,1996), felt emotion may help strengthen neural networks responsible for involuntary expressions, which in turn may enhance the mechanism of facial mimicry. Tracking the facial expressions and emotional experiences of participants during choir sessions may have provided some insight into the relationship between mimicry, the extent/duration of active expressiveness practice, and felt emotion.

Emotion Perception

Despite our predictions, no significant prechoir and postchoir differences were found on the emotion perception task for happy and sad emotions. Given our small sample size, however, greater power may have revealed an effect for the perception of happy emotions. Not only does the *p* value of .087 approach significance, but the effect size of .49 represents a medium-to-large effect, indicating that almost 25% of the variance between timepoints may be explained by the effect (Cohen, 1988; Cohen, 1992).

It is also possible that there was little room for improvement in accuracy for happy expressions, since it is generally easier to recognize than other emotions (due to the non-mimicry-related factor discussed earlier), particularly with a somewhat restrictive accuracy scale of 0-6. Happy expressions did in fact yield one of the highest prechoir mean accuracy scores of all the emotions, meaning a ceiling effect may have been present despite the use of more challenging stimuli. Importantly, these possible explanations mean that improved mimicry of happy emotions may correspond with improved emotion perception despite the non-significant finding.

In terms of the lack of improvement in the perception of sad emotions, this finding is not surprising in the context of the earlier discussions of why sad mimicry failed to improve. That is, an inadvertent lack of emphasis on producing sad facial expressions while singing (and/or lack of experiencing felt emotion as a result of singing sad songs) may have prevented significant improvements in mimicry of sad expressions and, in turn, significant improvements in the perception of sad emotions. Notably, if we had found a significant improvement in perception of sad emotions in the absence of corresponding improvements in facial mimicry, this interpretation would be undermined.

It may also be important to recall that several studies have found that medicated PD samples, even at moderate stages of disease progression, showed little impairment in the perception of most emotions (Sprengelmeyer et al., 2003). It is difficult to make inferences about the effects of medication or disease progression in our sample, since all were medicated and ranged in disease severity from moderate to severe. In addition, we found no relationship between stage of disease progression and our outcomes. Nonetheless, the fact that our PD sample was both medicated and required to be in an "on" state throughout testing (i.e., with the effects of their medication actively present) may mean that emotion perception ability was aided. The inclusion of an age-matched healthy control group may have been useful in determining whether there was in fact an emotion perception deficit in our PD group at baseline and, relatedly, whether healthy controls show relatively less benefit following choir participation. Alternatively, a comparison of performance between participants' "on" and "off" states in the PD group may also help isolate the effects of an active medicated state on performance on this task.

In consideration of the above discussion, the mimicry and emotion perception outcomes are interpretable in ways that are consistent with our hypotheses that choral singing can improve both mimicry and corresponding emotion perception. At least in the case of mimicry, they also suggest that active targeting of specific emotional expressions may be important.

Vocal Quality

From among our various vocal measures, significant improvements following choir participation were seen in lowest achievable pitch and jitter. In terms of pitch, it is interesting and somewhat surprising that while the lower boundary of pitch improved significantly and yielded a large effect size of $r = .58$ (Cohen, 1988; Cohen, 1992), the highest boundary of pitch did not only fail to improve significantly, but failed to increase at all following choir

participation. It is possible that because the pitch range task required participants to begin at their lowest pitch, they had run out of breath before reaching their true highest pitch, hence revealing fewer effects of practice. Another explanation may lie in the comfort range within which most individuals choose to sing. Moore (1991) and Killian and Buckner (2008) showed that both musically trained and untrained adults are most comfortable singing closer to the lower quartile of their pitch range. Since choir participants could choose their most comfortable octaves for singing, it is possible that most singers spent the majority of their singing time at the lower end of their natural pitch range, thus resulting in the likelihood of greater strengthening of the vocal apparatus in ways that help execute lower pitch. During informal observation of choir sessions, it was indeed noted that pitch comfort was frequently acknowledged by the choir director, which resulted in key signatures occasionally being adjusted to suit the majority of participants' comfort levels. As might be expected, this often led to lowering the octaves in which participants were required to sing, and again favouring exercise of the lower portion of the pitch range.

As predicted, jitter levels during the phonation duration task decreased following choir participation. It is notable that mean jitter levels in our sample were similar in elevation at prechoir to Tanaka et al.'s PD sample (2011) and decreased to just below the levels of their age-matched healthy control group at postchoir, suggesting restoration to normal levels. In perceptual terms, this change translates to a clearer, less harsh or coarse speaking voice. Conversely, though shimmer levels did decrease following choir participation, the pre-postchoir differences fell short of reaching significance ($p = .071$). It is possible, however, that a significant result would have been achieved with greater overall power, particularly considering the medium-to-large effect size that was yielded ($r = .37$) (Cohen, 1988; Cohen, 1992) and that was comparable to the effect size for improvements in jitter levels ($r = .42$).

We also failed to find any significant improvement in loudness following choir participation. Furthermore, the effect size for this measure was remarkably small ($r = .03$), suggesting that the non-significant finding is likely due to reasons other than lack of power. Problems with loudness in PD are often targeted by speech therapies, since a soft speaking voice leads to considerable communication difficulties (Fox et al., 2006). However, while speech therapies like LSVT LOUD are shown to have positive effects, studies investigating the effects of choral-based interventions have produced equivocal findings. For example, both Evans and colleagues (2012) and Yinger and Lapointe (2012) found improvements in loudness following a two-year fortnightly choir program and a six-week biweekly choir program respectively. However, an eight-week biweekly choral singing intervention investigated by Stegemöller et al. (2016) revealed no improvements in loudness. Furthermore, Shih and colleagues (2012) also failed to find any improvements in loudness as a result of a 12-week choir.

These differences in findings may be explained by the method with which loudness was measured. Importantly, only Evans et al. (2012) and Yinger and Lapointe (2012) measured loudness in conversational speech, whereas all other cited studies (our own included) utilized maximum loudness exercises for this measure (i.e., "say 'ah' as loudly as you can"). It is quite possible that these two exercises produce different results. In addition, while all the choirs in these studies included a range of vocal exercises, Yinger and Lapointe's (2012) focused specifically on loudness, with participants repeatedly cued to pay attention to and maximize their loudness while singing songs. While the choir director in our study encouraged louder singing on occasion, it was not an explicit instruction for session content. There is in fact evidence that self-perception of loudness and vocal vigilance play important roles in the mechanism underlying the impairment (Ho et al., 1999; Sapir et al., 2011; Shih et al., 2012). It is therefore likely that

improvements in loudness are facilitated by a consistent focus on aspects of this vocal feature in any intervention.

Although no significant improvement was observed in phonation duration following choir participation, this vocal measure yielded the strongest non-significant trend in the predicted direction and yielded a medium-to-large effect size of .38 (Cohen, 1988; Cohen 1992). The increase from prechoir to postchoir was close to what was observed in Stegemöller et al.'s study (2016), which did find significant improvement in phonation duration after eight weekly sessions of choral singing. Indeed, our *p* value of .06 for this measure could be seen as approaching significance, and possibly would have done so with greater statistical power. As with much of the vocal data, however, there are mixed findings in this area, since another study observed no improvement after 12 weekly sessions (Shih et al., 2012). Since singing often requires extending vocalizations beyond what is normal in speech in order to reach natural end-points indicated by lyrical content of the song, it is reasonable to expect improvements in phonation duration as a result of choir participation. However, as with loudness, this was not an explicit focus of our choir program and was not formally monitored. In addition, although the choir director occasionally encouraged extending an ending note and indicated where breaths should be taken, additional breaths were also sometimes recommended when sustaining a single breath posed as challenging. Although extent of focus on phonation duration in singing exercises may be an important factor, its effects so far remain unexplored to our knowledge.

It is difficult to reconcile the vocal measure outcomes of this study based on similar literature given the great variability in results. There may be multiple factors responsible for these differences, some of which have already been discussed. Outcomes may be influenced by differences in treatment duration, sample size, varying medication dosages, differences in testing

methods, differences in program content, or, perhaps more likely, a combination of several of these factors. It will take considerable further research to disentangle the roles of potentially important variables.

Limitations

Aside from issues with power, perhaps the most significant limitation of this study, as alluded to previously, is the absence of control groups. Without a comparison group of age-matched neurotypicals, we are unable to establish the presence and extent of impairment at baseline on our various measures. Similarly, an inactive control group of PD patients would have helped control for any effects of natural disease progression on our measures across the three-month study period, which may otherwise reduce the apparent magnitude of findings despite being clinically meaningful. As is often the case with repeated measures designs, extreme values within the dataset at one timepoint may have also resulted in regression toward the mean at the other timepoint and consequently led to erroneous conclusions. The inclusion of control groups would have helped address this limitation, as would the use of a various alternative designs. Unfortunately, due to time constraints of the study and a general difficulty in recruiting appropriate PD samples in the case of an inactive control group, we were not able to utilize control groups in this study.

Related to the issue of control, the lack of control or monitoring of how the choir program content was delivered may have affected our results and made subsequent interpretation challenging. The extent to which the choir director modeled and encouraged facial expressiveness, as well as the potentially disproportionate emphasis on particular songs, emotional expressions or vocal features may have affected participants' performance on our

outcome measures. Clearer instructions to the choir director and active tracking of delivered content may have been beneficial in this regard.

While we observed improvements on some of our measures, we are unable to determine how long-lasting these effects may be in the absence of follow-up testing. It is possible, for example, that maintenance of any improvements is dependent on continued engagement in a choral singing program, without which the positive effects may quickly diminish.

Another limitation of this study is the use of measures that lack well-established reliability. For example, the vocal tasks used to elicit the vocalizations from which the acoustic features were drawn – though commonly used – lack formally established reliability. As a result, it is possible that the degree of sensitivity of the measures differed across testing timepoints. For instance, although the loudness task ("Say 'ah' as loudly as you can") is often repeated in studies to increase the likelihood of capturing the participant's best effort, these instructions are nonetheless not always clear and may result in inconsistency between trials or between participants, such as yelling. Stegemöller et al. (2016), whose study also used this particular task, also pointed out this potential problem. Differences in reliability across different tasks may also be problematic, as they may adversely affect the pattern of results and lead to erroneous interpretations.

Finally, it is possible that participants who are current or former smokers may have had compromised vocal abilities as a result of smoking. However, we did not ask participants to report their smoking status. Similarly, a history of respiratory conditions (e.g., asthma) may have affected vocal abilities but were not recorded as part of the participant demographics.

Future Directions

Given that this study was reasonably exploratory in nature, there are many opportunities to gain further knowledge in the various areas of interest through follow-up research. For example, as noted previously, the effects of choral singing on outcomes relevant to this study have not yet been explicitly compared to individually-based singing, to our knowledge. Should such a comparison reveal equal benefits of the latter, it may inform the development of singing interventions for individuals with PD whose restricted mobility prevent them from attending choir sessions in the community.

Along the same lines, comparison groups involving different aspects of the intervention used in the present study may help determine which components contribute to the outcomes of interest. For example, it may be worthwhile comparing interventions with and without: (1) a social component; (2) a music-based component, and; (3) a music-based component with concurrent focus on facial/vocal muscles (rather than other motor functions). Such comparisons may help clarify the importance of the social and music-based aspects of such interventions, as well any potential benefit of focused versus unfocused motor activity in the presence of rhythmic auditory stimulation.

The present study investigated the impact of choral singing on the acoustic-based vocal impairments in PD rather than the more qualitative impairments such as speech intelligibility and articulation. It may be worthwhile utilizing professional raters such as speech-language pathologists in this endeavour, and also to determine possible improvements in effectiveness speech in contexts more true-to-life than brief vocal tasks.

Finally, since the ultimate aim of rehabilitating emotional communication impairments is to improve the quality of life of people living with PD, research to this end is equally important.

Some studies have already explored such benefits of choir participation (e.g., Pacchetti et al., 2000; Stegemöller et al., 2016), which can be complemented with the investigation of effects on broader areas of psychosocial functioning. For example, it would be useful to explore whether choir participation reduces psychosocial impairments such as depression and anxiety, both of which are common in PD populations. Anecdotally, it was clear that participants in our choir highly enjoyed participating in the program, and many reported subjective benefits in not just communication ability, but also confidence and overall engagement with others. In fact, the choir was so popular that it inspired the establishment of an independent program with a permanent home at the Canadian Opera Company and is currently in the process of being registered as a charity. Clearly, such programs have the potential to bear many fruit for this population.

Conclusion

Effective emotion communication is essential to human life, but is impaired in several ways in PD. Based on marginally significant data, our study demonstrated that singing happy songs with corresponding happy facial expressions in a 13-week choir program may improve the facial mimicry of happy expressions, which are shown to be impaired in PD. Though less clear, this may translate to improvements in the perception of happy expressions in others, which also appear to be impaired in PD. Since no improvements were seen in mimicry or perception of emotions not specifically targeted by the choir program, it is likely that improvement is contingent on specificity. Finally, choir participation was found to improve lowest achievable pitch and jitter levels. These benefits are among the many that choral singing can potentially offer PD populations, with other potential benefits worthy of research including improvements in general wellbeing and psychosocial functioning. Importantly, choral singing is a treatment

format that is both cost effective and engaging, and thus likely to encourage continued

participation.

References

Argaud, S., Delplanque, S., Houvenaghel, J. F., Auffret, M., Duprez, J., Vérin, M., ... & Sauleau, P. (2016). Does Facial Amimia Impact the Recognition of Facial Emotions? An EMG Study in Parkinson's Disease. *PloS One*, *11*.

Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. F. (1996). Comparison of Beck Depression Inventories-IA and-II in psychiatric outpatients. *Journal of Personality Assessment*, *67*, 588-597.

Biehl, M. (1997). Matsumoto and Ekman's Japanese and Caucasian facial expressions of emotion (JACFEE): Reliability data and cross-national differences. *Journal of Nonverbal Behavior*, *21*, 3-21.

Buxton, S. L., McDonald, L., & Tippett, L. J. (2013). Impaired recognition of prosody and subtle emotional facial expressions in Parkinson's disease. *Behavioural Neuroscience, 127*, 193–203.

Calder, A. J., Keane, J., Manly, T., Sprengelmeyer, R., Scott, S., Nimmo-Smith, I., & Young, A. W. (2003). Facial expression recognition across the adult life span. *Neuropsychologia, 41*, 195-202.

Calder, A. J., Lawrence, A. D., & Young, A. W. (2001). Neuropsychology of fear and loathing. *Nature Reviews: Neuroscience, 2*, 352-363.

Calvo, M. G., & Lundqvist, D. (2008). Facial expressions of emotion (KDEF): Identification under different display-duration conditions. *Behavior Research Methods*, *40*, 109-115.

Cannito, M. P., Suiter, D. M., Beverly, D., Chorna, L., Wolf, T., & Pfeiffer, R. M. (2012). Sentence intelligibility before and after voice treatment in speakers with idiopathic Parkinson's disease. *Journal of Voice*, *26*, 214-219.

Colcher, A., & Simuni, T. (1999). Clinical manifestations of Parkinson's disease. *Medical Clinics of North America*, *83*, 327–347.

Cohen, J. (1988). Statistical power analysis for the behavioural sciences. Hillside. *NJ: Lawrence Earlbaum Associates*.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155-159.

Cummings, J. L. (1992). Depression and Parkinson's disease: A review. *The American Journal of Psychiatry, 149*, 443-54.

Dalili, M. N., Penton-Voak, I. S., Harmer, C. J., & Munafò, M. R. (2015). Meta-analysis of emotion recognition deficits in major depressive disorder. *Psychological Medicine*, *45*, 1135-1144.

de Bruin, N., Doan, J. B., Turnbull, G., Suchowersky, O., Bonfield, S., Hu, B., & Brown, L. A. (2010). Walking with music is a safe and viable tool for gait training in Parkinson's disease: The effect of a 13-week feasibility study on single and dual task walking. *Parkinson's Disease*, *2010*.

Desmurget, M., Grafton, S.T., Vindras, P., Gre ́a, H., & Turner, R.S. (2004). The basal ganglia network mediates the planning of movement amplitude. European Journal of Neuroscience, 19, 2871–2880.

Dimberg, U., & Petterson, M. (2000). Facial reactions to happy and angry facial expressions: Evidence for right hemisphere dominance. *Psychophysiology*, *37*, 693–696.

Dumer, A. I., Oster, H., McCabe, D., Rabin, L. A., Spielman, J. L., Ramig, L. O., & Borod, J. C. (2014). Effects of the Lee Silverman Voice Treatment (LSVT® LOUD) on Hypomimia in Parkinson's Disease. *Journal of the International Neuropsychological Society*, *20*, 302-312.

Dujardin, K., Blairy, S., Defebvre, L., Duhem, S., Noel, Y., Hess, U., & Destée, A. (2004). Deficits in decoding emotional facial expressions in Parkinson's disease. *Neuropsychologia*, *42*, 239-250.

Ekman, P. (1992). Are there basic emotions? *Psychological Review, 99*, 550-553.

Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). The Duchenne smile: Emotional expression and brain physiology: II. *Journal of Personality and Social Psychology*, *58*, 343-353.

Ekman, P., & Friesen, W. V. (1976). Measuring facial movement. *Environmental Psychology and Nonverbal Behavior*, *1*, 56-75.

El Sharkawi, A., Ramig, L., Logemann, J., Pauloski, B., Rademaker, A., Smith, C., y Werner, C. (2002). Swallowing and voice effects of Lee Silverman Voice Treatment: A pilot study. *Journal of Neurology, Neurosurgery, and Psychiatry, 72*, 31–36.

Elefant, C., Baker, F. A., Lotan, M., Lagesen, S. K., & Skeie, G. O. (2012). The effect of group music therapy on mood, speech, and singing in individuals with Parkinson's disease—A feasibility study. *Journal of Music Therapy*, *49*, 278-302.

Elefant, C., Lotan, M., Baker, F. A., & Skeie, G. O. (2012). Effects of music therapy on facial expression of individuals with Parkinson's disease: A pilot study. *Musicae Scientiae*, *16*, 392-400.

Evans, C., Canavan, M., Foy, C., Langford, R., & Proctor, R. (2012). Can group singing provide effective speech therapy for people with Parkinson's disease? *Arts & Health*, *4*, 83-95.

Ferrand, C. T. (2007). *Speech Science: An integrated approach to theory and clinical practice* (2nd ed.). Boston, MA: Allyn and Bacon.

Fogg-Rogers, L., Buetow, S., Talmage, A., McCann, C. M., Leão, S. H., Tippett, L., ... & Purdy,

    S. C. (2016). Choral singing therapy following stroke or Parkinson's disease: an

    exploration of participants' experiences. *Disability and Rehabilitation*, *38*, 952-962.

Fox, C. M., Ramig, L. O., Ciucci, M. R., Sapir, S., McFarland, D. H., & Farley, B. G. (2006).

    The science and practice of LSVT/LOUD: Neural plasticity-principled approach to

    treating individuals with Parkinson disease and other neurological disorders. *Seminars in*

    Speech and Language, *27*, 283-299.

Frank, M. G., Ekman, P., & Friesen, W. V. (1993). Behavioral markers and recognizability of the

    smile of enjoyment. *Journal of Personal and Social Psychology*, *64*, 83–93.

Gelb, D. J., Oliver, E., & Gilman, S. (1999). Diagnostic criteria for Parkinson's disease. *Archives*

    of Neurology, *56*, 33.

Goetz, C. G., Poewe, W., Rascol, O., Sampaio, C., Stebbins, G. T., Counsell, C., et al. (2004).

    Movement disorder society task force report on the Hoehn and Yahr staging scale: Status

    and recommendations the movement disorder society task force on rating scales for

    Parkinson's disease. *Movement Disorders, 19*, 1020–1028.

Hackney, M. E., Kantorovich, S., Levin, R., & Earhart, G. M. (2007). Effects of tango on

    functional mobility in Parkinson's disease: a preliminary study. *Journal of Neurologic*

    Physical Therapy, *31*, 173-179.

Harding, A. J., Stimson, E., Henderson, J. M., & Halliday, G. M. (2002). Clinical correlates of

    selective pathology in the amygdala of patients with Parkinson's disease. *Brain : A*

    Journal of Neurology, 125, 2431-2445.

Heida, T., Poppe, N. R., de Vos, C. C., van Putten, M. J. A. M., & van Vugt, J. P. P. (2014).

    Event-related mu-rhythm desynchronization during movement observation is impaired in

    Parkinson's disease. *Clinical Neurophysiology*, *125*, 1819-1825.

Herd, C. P., Tomlinson, C. L., Deane, K. H., Brady, M. C., Smith, C. H., Sackley, C. M., &

    Clarke, C. E. (2012). Comparison of speech and language therapy techniques for speech

    problems in Parkinson's disease. *The Cochrane Library*.

Hess, U., Blairy, S., & Kleck, R. E. (1997). The intensity of emotional facial expressions and

    decoding accuracy. *Journal of Nonverbal Behavior*, *21*, 241-257.

Ho, A.K., Bradshaw, J.L., Iansek, R., & Alfredson, R. (1999). Speech volume regulation in

    Parkinson's disease: Effects of implicit cues and explicit instructions. *Neuropsychologia*,

    37, 1453–1460.

Holmes, R. J., Oates, J. M., Phyland, D. J, & Hughes, A. J. (2000). Voice characteristics in the

    progression of Parkinson's disease. *International Journal of Language & Communication*

    Disorders, *35*, 407-418.

Jürgens, U. (2002). Neural pathways underlying vocal control. *Neuroscience and Biobehavioral*

    Reviews, 26, 235–258.

Kan, Y., Kawamura, M., Hasegawa, Y., Mochizuki, S., & Nakamura, K. (2002). Recognition of

    emotion from facial, prosodic and written verbal stimuli in Parkinson's disease. *Cortex,*

    38, 623-630.

Keane, J., Antoun, N., Manes, F., Calder, A. J., & Young, A. W. (2000). Impaired recognition

    and experience of disgust following brain injury. *Nature Neuroscience, 3*, 1077-1078.

Killian, J. N., & Buckner, J. J. (2008). Comparison of starting pitch preferences among fourth

    graders, undergraduate music majors, and elementary education majors. *Music Education*

Research International, 2, 11-20.

Ledger, S., Galvin, R., Lynch, D., & Stokes, E. K. (2008). A randomized controlled trial evaluating the effect of an individual auditory cueing device on freezing and gait speed in people with Parkinson's disease. *BMC Neurology*, *8*, 46.

Liotti, M., Ramig, L.O., Vogel, D., New, P., Cook, C.I., Ingham, R.J., & Fox, P.T. (2003). Hypophonia in Parkinson's disease: Neural correlates of voice treatment revealed by PET. *Neurology*, *60*, 432–440.

Livingstone, S. R., Peck, K., & Russo, F. A. (2012). *RAVDESS: The Ryerson Audio-Visual Database of Emotional Speech and Song*, in Canadian Society for Brain, Behavior and Cognitive Science: Kingston, ON.

Livingstone, S. R., Vezer, E., McGarry, L. M., Lang, A. E., & Russo, F. A. (2016). Deficits in the mimicry of facial expressions in Parkinson's disease. *Frontiers In Psychology*. doi:10.3389/fpsyg.2016.00780

Livingstone, S. R., Vezer, E., McGarry, L. M., Lang, A. E., & Russo, F. A. (in progress). A novel singing therapy for rehabilitating facial and vocal expressive deficits in Parkinson's disease.

Lundqvist, L. O. (1995). Facial EMG reactions to facial expressions: A case of facial emotional contagion? *Scandinavian Journal of Psychology, 36*, 130–141.

Marsili, L., Agostino, R., Bologna, M., Belvisi, D., Palma, A., Fabbrini, G., & Berardelli, A. (2014). Bradykinesia of posed smiling and voluntary movement of the lower face in Parkinson's disease. *Parkinsonism & Related Disorders*, *20*, 370-375.

McCade, D., Savage, G., & Naismith, S. L. (2012). Review of emotion recognition in mild cognitive impairment. *Dementia and Geriatric Cognitive Disorders, 32*, 257-266.

McDonald, W. M., Richard, I. H., & DeLong, M. R. (2003). Prevalence, etiology, and treatment of depression in Parkinson's disease. *Biological Psychiatry*, *54*, 363-375.

McIntosh, G. C., Thaut, M. H., Rice, R. R., & Miller, R. A. (1994). Stride frequency modulation in parkinsonian gait using rhythmic auditory stimulation. *Annals of Neurology*, *36*, 316-316.

Moore, R. S. (1991). Comparison of children and adults' vocal ranges and preferred tessituras in singing familiar songs. *Bulletin for the Council for Research in Music Education, 107*, 13-22.

Narayana, S., Fox, P.T., Zhang, W., Franklin, C., Robin, D.A., Vogel, D., & Ramig, L.O. (2010). Neural correlates of efficacy of voice therapy in Parkinson's disease identified by performance-correlation analysis. *Human Brain Mapping, 31*, 222–236.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., ... & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*, 695-699.

Neal, D. T., & Chartrand, T. L. (2011). Embodied emotion perception amplifying and dampening facial feedback modulates emotion perception accuracy. *Social Psychological and Personality Science*, *2*, 673-678.

Niedenthal, P. M., Brauer, M., Halberstadt, J. B., & Innes-Ker, Å. H. (2001). When did her smile drop? Facial mimicry and the influences of emotional state on the detection of change in emotional expression. *Cognition and Emotion*, *15*, 853-864.

Pacchetti, C., Mancini, F., Aglieri, R., Fundarò, C., Martignoni, E., & Nappi, G. (2000). Active

music therapy in Parkinson's disease: an integrative method for motor and emotional

rehabilitation. *Psychosomatic Medicine*, *62*, 386-393.

Pentland, B., Pitcairn, T. K., Gray, J. M., & Riddle, W. (1987). The effects of reduced expression

in Parkinson's disease on impression formation by health professionals. *Clinical

Rehabilitation*, 307-312.

Pinto, S., Ozsancak, C., Tripoliti, E., Thobois, S., Limousin-Dowsey, P., & Auzou, P. (2004).

Treatments for dysarthria in Parkinson's disease. *The Lancet Neurology*, *3*, 547-556.

Pitcairn, T. K., Clemie, S., Gray, J. M., & Pentland, B. (1990). Non-verbal cues in the self-

presentation of Parkinsonian patients. *British Journal of Clinical Psychology*, *29*, 177-

184.

Ramig, L. O., Sapir, S., Countryman, S., Pawlas, A. A., O'brien, C., Hoehn, M., & Thompson, L.

L. (2001). Intensive voice treatment (LSVT®) for patients with Parkinson's disease: a 2

year follow up. *Journal of Neurology, Neurosurgery & Psychiatry*, *71*, 493-498.

Rinn, W. E. (1984). The neuropsychology of facial expression: A review of the neurological and

psychological mechanisms for producing facial expressions. *Psychological Bulletin*, *95*,

52-77.

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review

of the cross-cultural studies. *Psychological Bulletin, 115*, 102-141.

Sapir, S., Ramig, L.O., & Fox, C.M. (2011). Intensive voice treatment in Parkinson's disease:

Lee Silverman Voice Treatment. *Expert Reviews in Neurotherapeutics, 11*, 815–830.

Sato, W., Kubota, Y., Okada, T., Murai, T., Yoshikawa, S., & Sengoku, A. (2002). Seeing happy

emotion in fearful and angry faces: Qualitative analysis of facial expression recognition

in a bilateral amygdala-damaged patient. *Cortex, 38*, 727-742.

Shih, L. C., Piel, J., Warren, A., Kraics, L., Silver, A., Vanderhorst, V., ... & Tarsy, D. (2012). Singing in groups for Parkinson's disease (SING-PD): A pilot study of group singing therapy for PD-related voice/speech disorders. *Parkinsonism & Related Disorders*, *18*, 548-552.

Shulman, L. M., Taback, R. L., Rabinstein, A. A., & Weiner, W. J. (2002). Non-recognition of depression and other non-motor symptoms in Parkinson's disease. *Parkinsonism & Related Disorders*, *8*, 193-197.

Simons, G., Ellgring, H., & Pasqualini, M. C. S. (2003). Disturbance of spontaneous and posed facial expressions in Parkinson's disease. *Cognition & Emotion*, *17*, 759-778.

Simons, G., Pasqualini, M. C. S., Reddy, V., & Wood, J. (2004). Emotional and nonemotional facial expressions in people with Parkinson's disease. *Journal of the International Neuropsychological Society*, *10*, 521-535.

Skodda, S. (2012). Effect of deep brain stimulation on speech performance in Parkinson's disease. *Parkinson's Disease*, *2012*.

Smith, M. C., Smith, M. K., & Ellgring, H. (1996). Spontaneous and posed facial expression in Parkinson's disease. *Journal of the International Neuropsychological Society*, *2*, 383–391.

Spielman, J.L., Borod, J.C., & Ramig, L.O. (2003). The effects of intensive voice treatment on facial expressiveness in Parkinson disease: Preliminary data. *Cognitive and Behavioral Neurology, 16*, 177–188.

Spielman, J. L., Dumer, A., Borod, J. C., Oster, J., Rabin, L. A., Halpern, A., & Ramig, L. O. (2012, February). Impact of intensive treatment targeting loudness or articulation on

facial and vocalexpression in Parkinson disease (PD). Paper presented at the 16th Biennial Conference on Motor Speech, Santa Rosa, CA.

Sprengelmeyer, R., Young, A. W., Mahn, K., Schroeder, U., Woitalla, D., Büttner, T., . . . Przuntek, H. (2003). Facial expression recognition in people with medicated and unmedicated Parkinson's disease. *Neuropsychologia, 41,* 1047–1057.

Suzuki, A., Hoshino, T., Shigemasu, K., & Kawamura, M. (2006). Disgust-specific impairment of facial expression recognition in Parkinson's disease. *Brain: A Journal of Neurology, 129,* 707–717.

Tanaka, Y., Nishio, M., & Niimi, S. (2011). Vocal acoustic characteristics of patients with Parkinson's disease. *Folia Phoniatrica et Logopaedica*, *63*, 223-230.

Teixeira, J. P., Oliveira, C., & Lopes, C. (2013). Vocal acoustic analysis–jitter, shimmer and HNR parameters. *Procedia Technology*, *9*, 1112-1122.

Thaut, M. H., Kenyon, G. P., Schauer, M. L., & McIntosh, G. C. (1999). The connection between rhythmicity and brain function. *IEEE Engineering in Medicine and Biology Magazine*, *18*, 101-108.

Thaut, M. H., McIntosh, G. C., Rice, R. R., Miller, R. A., Rathbun, J., & Brault, J. M. (1996). Rhythmic auditory stimulation in gait training for Parkinson's disease patients. *Movement Disorders*, *11*, 193-200.

Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience, 18*, 411.

Wild, B., Rodden, F.A., Grodd, W., & Ruch, W. (2003). Neural correlates of laugher and humor. *Brain, 126,* 2121–2138.

Winston, J. S., O'Doherty, J., & Dolan, R. J. (2003). Common and distinct neural responses during direct and incidental processing of multiple facial emotions. *Neuroimage, 20*, 84-97.

Yinger, O. S., & Lapointe, L. L. (2012). The Effects of Participation in a Group Music Therapy Voice Protocol (G-MTVP) on the Speech of Individuals with Parkinson's Disease. *Music Therapy Perspectives*, *30*, 25-31.

Zhou, R., & Hu, S. (2004). Effects of viewing pleasant and unpleasant photographs on facial EMG asymmetry. *Perceptual and Motor Skills*, *99*, 1157–1167.