The pre-publication draft version of "Origins and Functions of Clarinettists' Ancillary Gestures." Marcelo M. Wanderley and Bradley Vines. In A. Gritten and E. King, eds. *Music and Gesture: New Perspectives on Theory and Contemporary Practice*, Ashgate Publishing, 2006.

Ancillary Gestures of Clarinettists

Marcelo M. Wanderley and Bradley W. Vines

Musical performance is a revealing human behaviour for the analysis of communication due to its expressive content. Not only are traditional musical parameters conveyed – melody, rhythm, articulation – but also information about emotions experienced by the performer and intended by the composer. In combination, these variables contribute to an observer's overall experience. The analysis of musical performance is therefore a broad research subject based upon knowledge and methods from several domains (Gabrielsson 1999; 2003), which complements research on speech and gesture (McNeill 1992; 2000).

Concerning the study of performer gestures (Cadoz and Wanderley, 2000), as compared to the study of other aspects of musical performance, until recently researchers have paid little attention to the gestural behaviour of instrumentalists (Gabrielsson 1999). One early (and insightful) investigation on expressive movements of musicians was performed by François Delalande (1988, 1990). In an in-depth study of several videos filmed by Bruno Montsaingeon of Glenn Gould playing various pieces of J. S. Bach, Delalande (1988) made several observations concerning the musical implications of Gould's accompanist gestures. For instance, in the beginning of the first fugue from J. S. Bach's *The Art of The Fugue* where only the right hand performs, Gould's left hand seemed to 'conduct an imaginary orchestra' (p. 86). According to Delalande (p. 92):

> It should then be concluded that at least in this example, Gould's behaviour is divided between two orientations. One, incarnated by the left hand, is a reading of the score while the other, directed towards the production of the sound object, is performed by the right hand.[1]

From the study of roughly one hour of Gould's performances, Delalande proposed a typology of gestures in three levels, from a purely functional to a purely symbolic one:

1. *Effective gestures*, those that actually produce the sound;
2. *Accompanist gestures*, associated body movements;

3. *Figurative gestures*, gestures perceived by a listener, but without a direct correspondence to a movement of the performer. Examples include changes in note articulation, melodic variations, etc.

Accompanist gestures include postures and movements of the head, chest, (left) hand (when only the right hand performs), shoulders, elbows and eyebrows. These gestures are specific configurations of positions, movements, and imitations and are of five types:

1. *Recueilli* (*meditative, contemplative*): chest immobile and leaning forward, lowered head, chin almost touching the chest; occasional eyebrow movements; vertical movement of left hand if free;
2. *Vibrant* (*vibrant*): chest curves and straightens up alternatively, lower back immobile; occasional head, eyebrow and various left hand movements;
3. *Fluant* (*fluid*): chest oscillates forward and backward or turns, occasional eyebrow movements;
4. *Délicat* (*delicate*): chest is immobile, leaning markedly forward, nose almost touching the keyboard;
5. *Vigoureux* (*vigorous*): chest is immobile, vertical and a bit forward, head close to the shoulders, shoulders forward.

In addition, Delalande analysed the eyebrows more extensively. He identified four types: *short eyebrow raises* (with meditative and fluid gestures); *eyebrow raises accompanying melodic motives* (with vibrant gestures); *long eyebrow raises* (with delicate gesture) and *knitted eyebrows* (with vigorous gestures). Delalande observed that the above cases happen as 'pure cases' (frequently, with clear transitions between the types) and 'impure' cases (more rarely, when they are not clearly defined).

The above gesture types (or *styles*) are related to a *temporal segmentation* of the score into units of different sizes. For instance, the *delicate style* is related to autonomous notes; the *vibrant style* to short, well-defined melodic motifs; the *fluid style* to long, continuous phrases; the *vigorous style* being less associated with timing, but more often with marked chords or rhythms. Among the various other findings in this work, the most interesting is perhaps that Gould's movements, although related to the score, did not correspond one-to-one with the notes on the score or to orchestral imaging, but were clearly linked to an *emotional content* in his performance.

In 1993, Jane Davidson published a study on the perception of visual cues from expressive movements of four violin players as well as of a pianist. She showed that visual information about the musicians' body movements conveyed the intended expressive performance manner (standard, exaggerated or deadpan) more clearly than the combined video/sound or sound alone presentation conditions. Davidson's results showed that performance gestures are rich with information

about expressive intention. We return to Davidson's research in our discussion of the psychological impact of musicians' expressive movements.

Both Delalande's and Davidson's studies have shown that musicians not only perform skilled movements that are directly related to sound production, but also movements that do not seem to have an obvious link to the generation of sound. Furthermore, Davidson showed that these movements convey meaning to the musician and to the audience. These works have focused on piano (Delalande 1988), and piano and violin (Davidson 1993). This chapter contains an overview of some of our experimental work on ancillary gestures of clarinettists. We explore these gestures from two perspectives: first, by analysing the role of expressive movements in music performance; and second, by revealing the influence of such gestures on audience members' perception of music.

1.1 Quantitative Analysis of Ancillary Gestures

This research focuses on ancillary gestures of clarinet players as they perform selections from the clarinet repertoire. The terms *ancillary* or *accompanist gestures* have been used to designate those gestures that are part of a performance, but not produced in order to generate sound. These terms have a similar meaning to what Davidson (1993) called *expressive movements*, although they are to be differentiated from *expressive gestures* in the sense presented by Orio (1999). To elaborate, one way to classify gestures is by analysing subtle variations of a basic gesture (whether 'manipulative' – in contact with an object – or 'empty-handed') that communicate *expressive content* (Camurri, Mazzarino, Ricchetti, Timmers and Volpe 2004). The main objective is to separate *basic* gestures that are part of a vocabulary of pre-established actions from *variations* on these basic gestures. Authors have used this approach to distinguish between *symbolic* and *parametric* gestures (Modler and Zannos 1997) or *gestures* and *gesture nuances* (Orio 1999).

Orio used an approach based on the expressive content of a movement to study the gestures of an acoustic guitarist. He considered the position of left-hand fingers on the strings (defining pitch) and right-hand string-picking pressure (defining loudness) as gestures belonging to defined *classes* - pitch and loudness – that comprise the basic level of information transmitted to the audience. The way instrumentalists perform these two actions produces a second level of information, that of *expressive content*, or *gesture nuance*. Orio related this second level to the information contained in the instrument's timbre.

In the context of clarinet performance, ancillary gestures are a class of wind instrument performer gestures that are produced by moving the instrument during performance - *lifting up/lowering down*, *to one side or the other*, *fast tilt-like gestures*, *head, waist, knee movements*, and *associated body postures* (Wanderley 1999). Although ancillary gestures may be produced consciously – in accordance with a composer's explicit requirements (as in some pieces by Karlheinz Stockhausen and Anton Webern), as a visual effect for the audience (Cook 2003; Davidson 2001), or as part of a communication language between players in an ensemble (Williamom and Davidson 2002) – our goal is to analyse musicians'

performance gestures *that do not seem to have a well-defined purpose* related to sound production, inter-performer, or symbolic performer–audience communication.

We have collected quantitative measurements of several clarinettists playing pieces from the standard clarinet repertoire (Wanderley 2002; Wanderley, Vines, Middleton, McKay and Hatch 2005). Excerpts of the following recorded pieces have been analysed as part of this research: *Domaines* by Pierre Boulez; *Sonata No. 1 for Clarinet and Piano* by Johannes Brahms (first movement); *Sonata for Clarinet* by Francis Poulenc (first and second movements); *Three Clarinet Pieces* by Igor Stravinsky (first and second pieces); and other works by Stockhausen, Mefano, and Gaussain.

Movement data were acquired with digital video cameras and a high-accuracy movement tracker (Optotrak 3020 infrared system[2]). Acquisition sessions took place at the NICI, Nijmegen, at the Free University Amsterdam, and at the Motor Control Laboratory, McGill University. Eight to ten active infrared markers were placed on the performer and clarinet. Performers stood about 3.5 metres away from the cameras. The Optotrak recorded a three-dimensional coordinate for each marker's position in space (horizontal, vertical and sagittal coordinates) at 100 Hz.

In this chapter, we will limit ourselves to the observation of the vertical movement of a marker placed on the clarinet bell. More holistic approaches, including the use of Laban-Bartenieff movement fundamentals, are currently being explored in our laboratory (Campbell Chagnon and Wanderley, in preparation).

Insert Figure 2 about here

We have focused on three primary question areas:

1. The *production* of ancillary gestures - whether it is common to move the instrument while playing, the size of these movements, identifying basic movement patterns, and the influence of skill level and other external factors on movement production.
2. The *repeatability* of ancillary gestures - whether a clarinettist repeats the same sequence of movements while playing a piece multiple times, and how movement patterns change with practice and over long time periods.
3. The *production of similar movement patterns by different performers*, i.e., whether there exist any movement patterns common to different performers and whether there exists a consistent relation between movements of different performers and the musical score.

In the following sections we will present evidence that illuminates answers to some of these questions (for a thorough analysis, see Wanderley 1999, 2002; Wanderley et al. 2005).

1.2 Production and Repeatability

It seems clear that most musicians produce ancillary gestures of various types while performing a piece.

Figure 4 shows the vertical movement of the clarinet bell as a musician performed part of the first movement of Poulenc's Clarinet Sonata. It should be noted that ancillary gestures were produced continually throughout the performance, with a maximum vertical movement of the bell in a range of approximately 22 centimetres.

Insert Figure 4 about here

1.2.1 Rocking Movements

Several performances involved 'rocking' movements (Davidson and Correia 2002), although not all of them. Rocking movements involved a wave-like oscillation of the height of the instrument, upper body and/or the knees (Wanderley et al. 2005). These movements were often strongly correlated with structural characteristics of the piece being performed (Wanderley 2002).

Insert Figure 5 about here

Four different performers made rocking movements while performing the first movement of the Brahms Sonata, although these differed across performers. In this case, the presence of rocking movements can be linked to the clear rhythmic nature of the music. This may not be the case for other pieces such as *Domaines* or the *Three Pieces for Solo Clarinet* that have no underlying pulse or metre to which the clarinettists might entrain their movements (Wanderley et al. 2005).

Vertical movement range of the clarinet bell varied across performers, the different heights of the performers may account for this variation.

The temporal segmentation of the movements was revealing as well: Performer 1 made longer movements (in duration) than the other three performers. Performer 1 moved up and down with roughly half the frequency of Performer 2 or Performer 3. Performer 4 used a mixed movement strategy, sometimes more similar to that of Performer 1, sometimes to that of the other two. Is Performer 1's unique movement pattern indicative of a different interpretation of the piece? This would be expected in light of evidence that body postures and movements reveal mental states, intentions and attitudes (Dittrich, Troscianko, Lea and Morgan 1996; McNeil 1992, 2000; Runeson and Frykholm 1983).

There were other consistencies for all four performers, such as the relative height of their bells during breathing, although this may not hold true for other pieces in which breathing is not as well defined as in this movement. It is nevertheless an

indication of the influence of respiration on the final shape of the movement contour that is largely constant across performers.

1.2.2 Movement Standing and Seated

Although it may seem unnatural to play some pieces seated – usually solo and sonata pieces are played in a standing position – during rehearsals clarinettists may play these pieces seated. Wanderley (2002) explored the influence of position (standing up or seated) on movements.

Insert Figure 7 about here

In the seated performance, the bell tip marker of interest was often obscured by the player's leg. However, it is clear that the bell movements for the seated position are similar to the movements for the standing position. Thus, the gesture sequence ingrained in a musician's memory for a piece is robust across position of performance.

1.2.3 Standard, Expressive and Immobilised

It is also interesting to analyse the influence of intended performance manner on the movements performed. We have asked performers to play in a *standard* manner (i.e. as in a recital), *expressively* (i.e. trying to exaggerate the *emotion* in the performance, but without making reference to a particular form of movement), and *immobilised* (i.e., trying *not to move* while playing). Note that the immobilized manner is different form Davidson's *deadpan* manner (Davidson, 1993), which specified an inexpressive performance with no explicit instruction to restrict movement. (Note, however, that Davidson did find a reduction of movement magnitude associated with the immobile performance manner; Davidson, 1994) We chose to use the immobile manner to determine whether it would be possible to play a piece accurately without expressive movements.

By analysing the data obtained with the Optotrak, one can carefully compare the movements associated with each of the three performance manners. Figure 10 displays data from three performances of Boulez's *Domaines*, Cahier A original (from top to bottom: standard, expressive and immobilised), plotted using a consistent vertical scale. The movement amplitudes are comparable for the first two performances, but are very different for the immobilised manner. This fact indicates that the Performer 2as indeed able to consciously suppress most ancillary gestures during performance, or, as we will see in Figure 11, to reduce the amplitude of the movements.

Insert Figure 10 about here

Insert Figure 11 about here

Other performers were also able to suppress movements[3], across performances of a variety of pieces. Note that this does *not* indicate that performers were *comfortable* playing the pieces while trying not to move. For instance, one performer stressed that in the immobilised manner breathing felt unnatural. All that this data shows is that it is possible to play most pieces while trying to suppress ancillary movements. In other words, ancillary movements are *not strictly necessary* for sound production.

In our experiments so far, not all performers made markedly larger movements in their expressive performance as compared to their standard performance, although this may happen in some cases, as shown in Figure 13.

Insert Figure 12 about here

Insert Figure 13 about here

1.3 Accurate Performance Comparisons

Although timing was very consistent for individual performances across repetitions of the same segment of music (considering that they lasted from 30 to around 80 seconds each), in order to accurately compare multiple performances from one or several performers, it was necessary to account for fluctuations in tempo throughout the performances and to correct for differences in overall duration. *Dynamic Time Warping* (or *data registering*) of performances facilitates an accurate comparison of the spatial movement characteristics for various performances by warping the movement data to a reference score (in this case, a midi file), thus eliminating timing fluctuations. The process involves first finding events (notes) in the audio signals available from the videos recorded in the data acquisition sessions. These event lists are then compared to the reference score – a perfectly quantized score – and the performances are adjusted to the reference timing. Using this method, the movements can be plotted with reference to a common timing.

Insert Figure 14 about here

Insert Figure 15 about here

Insert Figure 16 about here

Comparing two performers, one can see a general trend in their rocking movements, but several differences exist. More research is needed to fully understand the similarities and differences across performers.

1.3.1 Some Musical Considerations

We reported in the section on rocking movement that some performers tended to segment their movements according to different musical considerations (Wanderley et al. 2005). One can visualize this in the following figure that shows the bell movement of two performers playing the same excerpt of the Second Piece by Stravinsky. Note the different segmentation scheme used by the two performers.

Insert Figure 17 about here

1.4 Summary: Ancillary Gestures at Different Movement Levels

Based upon the findings presented in this chapter and in our previous work on ancillary gestures, we propose that these movements are related to musical features at different structural levels. From the various analyses presented in this chapter, we conclude that at least three levels influence clarinettists' expressive movements:

1. *Material/Physiological* - the influence of respiration, fingering, ergonomics of the instrument, etc.

2. *Rhythmic/Structural* - dependence on the characteristics of the piece being performed. We have seen that some performers have a tendency to mark the rhythm with their instrument in various ways (rocking movements). Although differences exist, there seems to be similarities that cannot be explained by randomness alone.

3. *Interpretative* - relations to the moment of interpretation of the piece as developed by the performer. These will likely be different for different performers.

But what do ancillary gestures communicate to an audience? What effect do they have on the perception of music performance? We will analyse these questions in the following section.

2.1 The Perceptual Significance of Clarinettists' Ancillary Gestures

Research by Wanderley and colleagues has focused upon the physical movements of clarinettists, with a particular focus on the performers' knowledge about their own movements and on the relationship between movement trajectories and musical interpretation. In addition, the sonic effects of movement gestures have been studied in detail, with respect to real-time sound synthesis and realness of sound (Wanderley and Depalle, 2004; Wanderley Depalle and Warusfel, 1999). In this section of the chapter, we focus on the psychological dimension of clarinettists' movements, including their effect on the audience's perception of a musical piece.

Davidson (1993), as mentioned above, empirically established the importance of musicians' movements in her study of expressive perception. That research showed that the visual aspect of a musical performance not only carries important information about the music and the musician's musical intentions, but may also convey information that augments the experience of sound alone.

Krumhansl and Schenck (1997) conducted seminal work by investigating emotional and structural responses to a ballet performance. They used a similar multi-modal approach to Davidson's (some participants only heard the music, some only saw the dance and the remainder both heard the music and saw the dance), however, Krumhansl and Schenck collected real-time judgments to gauge the experience of their participants *while* the performance was presented. They found that the dance conveyed much of the same emotional and structural information as did the music. This finding shows that the perception of movement and the perception of sound *can* create like experiences. It is possible that the movements of the dancers and the music of the orchestra accessed the same internal schemata, which Daniel Stern (1999) has referred to as 'vital contours'. Ballet is an art form in which the visual aspect is carefully predetermined, whereas a musician's movements are generally not choreographed (though some musicians may use movements to purposefully elicit a response in the audience, as mentioned above; Cook 2003; Davidson, 2001). Do the movements of musicians convey emotional and structural meaning that is similar to the experience of sound alone? The investigation discussed below addressed this very question.

We have pursued research to determine how musicians' movements influence an observer's perception of emotion and structure in musical performances (Vines, Wanderley, Krumhansl, Nuzzo and Levitin, 2004; Vines Krumhansl Wanderley Levitin, 2005; Vines Nuzzo and Levitin, in press;). The research has utilized a multi-modal approach, as did Davidson (1993) and Krumhansl and Schenck (1997), along with continuous measurements to determine the real-time emotional impact and perceptual significance of 'seeing' a musician perform. The experiments were designed to complement the work of Wanderley and colleagues, discussed above. Video recordings that were previously analyzed for their movement content (Wanderley 2002) were used as stimuli to explore the perceptual significance of that visual information for observers.

The following methodology was employed: Thirty musically trained participants saw, heard, or both saw and heard audio-video recordings of professional clarinettists performing Stravinsky's second piece for solo clarinet. This pattern of stimulus presentation is known as 'masking' in the analysis of auditory-visual interaction in cinema (Chion1994).

The participants made continuous judgments while observing the performances. They moved a slider up and down along a track seven centimetres in length to register their experience as time passed. A computer program recorded the slider location at 10 Hz and presented the stimuli simultaneously. The tasks were a continuous judgment of tension and a continuous judgment of phrasing. (The ordering of the two judgments was counterbalanced across participants so that half of them completed the phrasing judgment first.) For the tension judgment, participants were asked to express the tension they experienced in the performance by moving the slider upward as the tension increased and downward as the tension decreased. For the phrasing judgment, participants were instructed to move the slider upward as a phrase was entered and downward as a phrase was exited, so that the slider was near the top in the middle of a phrase and near the bottom between phrases.

The tension judgment has been found in past research to be a consistent measure that is responsive to many aspects of a musical stimulus (Fredrickson 1995, 2000; Krumhansl 1996). Krumhansl and colleagues showed that continuous judgments of tension were correlated with measures of affect and with a variety of physiological measures (Krumhansl 1997; Krumhansl and Schenck 1997). Tension in vision has not received as much attention in research, but dance theorists hypothesized that the following aspects of visual stimulation are likely to cause tension for an observer: initiation of movement, force, weight, rate of travel and rate of movement (Frego 1999), as well as the interaction between different forces, including those generated by the body and by the force of gravity (Sheets 1966).

The phrasing judgment revealed time points of perceptual segmentation for each participant's experience, as well as the sense of phrasing contour over time. Music, like language, organizes unique perceptual units into a hierarchical structure with meaning that unfolds over time (Cooper and Meyer 1960; Lerdahl and Jackendoff 1983; Levitin and Menon 2003). The musical phrase has an objective perceptual reality, as does the phrase in speech (Fodor and Bever 1965; Gregory 1978), and past research has used measures related to phrasing to study segmentation in the perception of music (Clarke and Krumhansl 1990; Deliège and El Ahmadi 1990; Krumhansl 1996; Krumhansl and Schenck 1997).

Techniques in the field of functional data analysis (Levitin Nuzzo Vines and Ramsay, 2005; Ramsay and Silverman 1997) were used to analyse the continuous measurements of tension and phrasing. These tools are ideal for use with judgments drawn from continuous processes, such as the experience of listening to a piece of music as it unfolds over time. Vines and colleagues have generated applications for functional data analysis in multi-modal research with music, and conceptual tools for research in musical emotion and temporal dynamics (McAdams Vines Vieillard Smith and Reynolds, 2004; Vines et al., 2005, in press).

Our data revealed that the visual aspect of the performances proved to contain a great deal of structural and emotional information. The clarinettists' movements, including their facial expressions, postures, breathing, and effective gestures, augmented participants' experience in three ways: 1) by reinforcing the information available in sound, 2) by contributing unique information to the overall experience, and 3) by conveying the performer's musical interpretation of the score.

The clarinettists that we chose for this research performed with very different movement styles. One performer, whom we will refer to as Performer 1, moved with smooth and controlled contours, for the most part. His feet never shifted during the performance and he rarely deviated from a front-facing orientation. The other performer, Performer 2, used highly idiosyncratic and expressive movements throughout the piece. For example, he would remain nearly motionless while playing fast runs until just at the end of the run, when he would have a great burst of expressive movement. These different movement styles led to variation in the tension induced while watching the performances, which we discuss below. Relations between audition and vision differed for the phrasing judgment and the tension judgment; hence we will consider the results of each measure in turn.

2.2 Tension Judgments

The visual aspect elicited different experiences of tension for each performer (see Figure 18). For those who could only see the performances, the mean response to Performer 1 was relatively static and low in magnitude. Visual only responses to Performer 2 were more dynamic and greater in overall magnitude. This is evidence that the impact of seeing a performance varies significantly across musicians. Different movement styles did generate different experiences of tension for the observers.

Insert Figure 18 about here

When both vision and audition were available to participants, the auditory component of the clarinet performances dominated the contour of experienced tension. Generally, the group mean for those who could both hear and see the performances was correlated with the group mean for those who could only hear the performances. The visual only contour followed a unique path for the most part; visual information generally conveyed a different dynamic contour of experienced tension compared to auditory information. There were some periods of time when the tension experienced from vision alone followed the same contour as the tension experienced by the auditory only group. The presence of such convergent sections, though few in number, confirms Krumhansl and Schenck's (1997) finding that visual stimuli and auditory stimuli can induce similar dynamics of emotion. Krumhansl and Schenck found a strong correlation between tension and emotion ratings elicited by seeing the dance without sound and by hearing the sound alone; the movement and music consistently elicited similar affective experiences in their study. Why were similar affective experiences for audition and vision relatively rare for the perception of musical performances? We posit that constraints due to playing the instrument restricted the musicians' ability to express emotion in visual terms. Dancers, such as those used in the ballet study, can move freely, whereas musicians have to maintain a consistent relationship with their instruments as they produce sound. The clarinet does allow for a variety of movements, which is one reason why we have chosen to focus on this instrument. However, the hand and arms are tightly constrained, as are facial expressions. Perhaps there would be greater similarity between emotion conveyed visually and emotion conveyed aurally for instruments that facilitate freedom of body movement and facial expression such as the piano, the guitar, and the voice. Future research will explore these other instruments.

Though sound largely determined the contour of tension experienced in the auditory+visual condition, visual input did affect the overall magnitude of tension experienced. This effect was especially pronounced when visual content contradicted auditory content - when the tension level in vision differed strongly

and consistently from the tension level in sound. For example, during the first section of Stravinsky's second piece for solo clarinet, Performer 1's calm and controlled movements contradicted the many high, loud and fast notes in sound. The tension registered by the auditory only group was very high while the tension registered by the visual only group was low. The auditory+visual mean followed the same contour as the auditory only group, but was lower during the section; that difference reached significance, as determined by a functional *F*-test. In this example, the visual information served to dampen the overall experience for those who could see as well as hear the performer. In another example, Performer 2's movements served to increase the overall experience of tension. His emotive and active movement pattern contradicted the quiet low notes in sound. The visual aspect induced a high magnitude of tension for those who could only see the performance while the tension induced by sound was low. Again, the auditory+visual group mean followed a contour that was similar to the auditory only mean, but with an increase in magnitude; that increase also reached significance by a functional *F*-test.

We also found evidence that the visual component can convey affective information in advance of the sound. An example of this kind of visual influence occurred for Performer 1. During a pause in sound, and before a new section began, Performer 1 adjusted his posture and facial expression (eye-brows rose perceptibly) to match the mood of the following section, which was lighter and lower in tension than the previous. The cues in gesture anticipated an affective change in the music and lead to a difference in perceived tension for participants who could both see and hear the performance as compared to those who could only hear it. The changes in Performer 1's expression gave those who could see the performance an indication of the emotional tone of the section to come, before the sound actually started. A functional data analysis (with phase-plane-plots) of the same transition segment for Performer 2 yielded a complementary finding (Vines et al., in press). The patterns of emotional experience conveyed by sound and by vision were similar in form, though the dynamics of change were shifted forward in time (occurred earlier) for participants who could see the performance. Thus, the movements of clarinettists can anticipate emotional changes in the music and can draw attention to them. Similar findings for phrasing are discussed below.

In general, the tension findings support the influence of a mirror system on communication between a performer and his or her audience. Such a neural system would enable emotional communication between performer and observer (Buccino Binkofski Fink Fadiga Fogassi Gallese et al. 2001; Rizzolatti and Arbib 1998). The movements and gestures did influence the experience of tension for observers, and the effect of seeing the musician depended upon idiosyncrasies in movement style.

2.3 Phrasing Judgments

Phrasing judgments yielded a different set of relations across sensory modalities than did tension ratings. We found evidence that a similar experience of structure was conveyed both by the visual aspect of the performances and by the sound

itself. Even though visual only participants did not hear the sound at all, they still extracted the underlying pattern of phrasing in the music. Many peaks and troughs in the mean curves were aligned for all presentation conditions, providing evidence for a similar pattern of perceptual segmentation across presentation conditions. This shows that the visual component of a musical performance can convey the structural information in a musical piece and that similar experiences of structural form can be conveyed through vision and hearing.

We chose to use the Stravinsky piece for our perception research because there is no underlying pulse or metre in the music; clarinettists tend to entrain their movements to the underlying pulse if there is one (Wanderley 2002). The players, therefore, were free to move about with idiosyncratic expressive patterns that were not constrained by any metrical consideration. In spite of these complex movement patterns, the participants who only saw the visual aspect of the performances were still able to abstract the structure of the music being played from the moving image. This implies that the movements of musicians are related to the musical structure, a finding that resonates with the work of Delalande (1988). It also shows that the human brain has an incredible capacity to abstract meaningful structure from complex movement patterns – though there was no simple relationship between gestures and musical phrasing, the viewers were able to recognize the underlying structure of the piece just by watching the clarinettist perform.

Further analyses showed that gestures played a role analogous to co-articulation movements in speech (Levelt 1989). The process known as *co-articulation* refers to 1) movements of vocal articulators that begin before a target sound is produced in speech and 2) the influence of previous muscular movements on the muscular pattern used to create the current syllable. In our musical context, the musicians' movements both anticipated the coming sound and extended beyond it; both kinds of movement influenced participants' judgment of phrasing in the piece. For example, certain gestures extended the sense of phrasing during a pause in sound. Though the sound came to an end at a major transition between musical sections, the performers' gestures and postures continued into the silence. As a result, those who could see the performers gauged a sense of phrasing that extended beyond the end of the note and into the silence. Similarly, certain gestures anticipated the beginning of the new section: Performer 2 swooped his clarinet downward before the new note began and Performer 1 took a breath and made a postural adjustment in anticipation of the coming sound. In this way, the clarinettists' movements served to cue the beginning of a new section for the participants. Clearly, musicians' gestures can complement the sound by anticipating the beginning of sound and by extending beyond the sound in different ways. An observer's experience of phrasing in a piece of music is enhanced by these extra-musical cues that create anticipation, expectation, and a sense of continuation that is complementary to the musical sound.

Body movements also served to convey the performers' musical interpretations of the score. An independent analysis of each performance by a music theorist revealed differences in the clarinettists' interpretations of phrasing in the Stravinsky piece. In particular, towards the end of the score, Performer 2 added an extra major phrase boundary, thus segmenting into two phrases what Performer 1

interpreted as one continuous phrase. Perceptual cues in both vision and sound conveyed the difference in performance intentions, which were mirrored in the participants' judgments across presentation conditions. This finding supports Runeson and Frykholm's (1983) assertion that body movements reveal mental states and intentions. In the case of musical performance, the intentions are related to intended perceptual segmentation and expressive content.

The visual modality was found to convey the underlying phrasing content in each performance, as well as the performers' individual musical intentions and interpretation of the score. The movement levels discussed above (Material/Physiological, Rhythmic/Structural, and Interpretative), and the constraints on movement that these levels entail, may account for the temporal correspondence between body movement and musical sound. However, in addition to providing redundant phrasing information with sound, musicians' body movements also serve to extend and to anticipate phrasing boundaries to influence the overall experience of an observer.

2.4 Summary: The Perception of Ancillary Gestures

In general, we have found that the visual aspect of musical performances contributes importantly at specific points in the music. The gestures and movements of clarinettists augment the experience of tension and the sense of phrasing by complementing, anticipating, following and sometimes contradicting the information available in sound. It is clear that the interaction between vision and sound in musical performance is rich and complex and that an observer's experience of the music is significantly enhanced when the performer can be seen as well as heard.

3.1 Conclusion

The study of gestures in music is an important research field for several reasons, including its contribution to psychological work on performance and cognition. In this chapter, we have focused on the analysis of ancillary gestures produced by clarinettists and their perception in performance. In the context of clarinet performance, our previous work has shown that ancillary gestures are common in performances, although not essential (clarinet players were eventually able to play excerpts of pieces with almost no expressive movements). For the case of the same expert clarinet player performing one piece at different times, a strong correlation between the player's movements at the same points in the score in the different performances was found, suggesting that ancillary gestures by clarinet players were not randomly produced, or just a visual effect, but that these gestures are an integral part of the performance process and mental representation of the music.

Quantitative data from performances of different players showed that ancillary gestures were idiosyncratic for each player, though movement features related to structural characteristics of the piece (e.g. tempo) and to material/physiological aspects tended to be largely invariant across performances of different musicians.

A three level typology of ancillary gestures according to their possible origin was proposed (*material/physiological*, *rhythmic/structural* and *interpretative*) in order to take into account similarities and differences among the different gestural patterns.

But apart from the analysis of pure movement, the study of ancillary gestures raises many questions about what a musician's movements convey to an audience. In every day speech, gestures of the hands and body, along with facial expressions, are continually reinforcing, modifying, or negating the meaningful content in speech (McNeil 2000). Non-speech gestures also contribute to the emotional essence of an utterance (Planalp DeFrancisco and Rutherford, 1996), and they help speakers to time their exchange in a conversation (McFarland, 2001). Do the expressive body movements of musicians play a similar role in relation to the musical sound? Music is analogous to speech in that the sound carries the core content.[4] Though a great deal of music is experienced without seeing the musicians (by means of CDs, MP3 players, computers and so on), many enthusiasts do seek opportunities to see live performances or to watch video recordings. We have found evidence that contributes to explaining why people enjoy seeing musical performances.

There may be a musical parallel to the proposition of Quek et al. (2002) that gesture and speech *proceed together from the same 'idea units.'* We posit that the musical sound and the musician's gestures (ancillary and effective) proceed from the same *performance 'idea unit.'* Such a relationship could explain the tight correspondence between expressive body movement and sound, as if part of the musical expression comes through the auditory sense, and the other part comes through the musician's gestures. This perspective is in accordance with Alexander Truslit's view that a musician's sound and body movements both originate in the same "inner motion" ("innere bewegtheit"; Repp, 1993). Future research will explore relations between musical sound and ancillary movements that are analogous to relations between vocal speech and paralinguistic gestures.

Acknowledgements

The authors would like to thank all of the performers for their patience and willingness to participate in the experiments, as well as the NICI, Nijmegen, the Free University Amsterdam and the Motor Control Laboratory, McGill University for the use of the Optotrak systems. Our warmest thanks to Daniel Levitin, Carol Krumhansl and Regina Nuzzo for their collaboration on various phases of this project. We are also grateful to other colleagues who have participated in this research, most notably Thomas Hélie, Nicola Orio, Cory McKay, Wesley Hatch and Neil Middleton. Research by the first author is funded through the Fonds Québecois de Recherche sur la Nature et les Technologies (FQRNT), the Natural Sciences and Engineering Research Council of Canada (NSERC) and Canadian Foundation for Innovation (CFI). The second author was partially funded by a J. W. McConnell McGill Majors Fellowship, and a CIRMMT Doctoral Fellowship.

Notes

1. 'On doit donc conclure qu'au moins dans cet exemple le comportement de Gould se partage entre deux orientations. L'une, incarnée par la main gauche, est une lecture de la partition alors que l'autre, tournée vers la réalisation de l'objet sonore, est concrétisée par le jeu de la main droite.'

2. Since July 2004, we have been using a Vicon System 460 with 6 M2 infrared cameras with passive markers, available at the Input Devices and Music Interaction Laboratory, McGill University.

3. Only once, one performer gave up one immobilised performance (Stravinsky, Second Piece) and claimed he could not do it. This was the only case in four performers and several pieces.

4. This statement might be less applicable to music traditions outside of the Western classical genre. Baily (1985), for example, argues that body movements are primary to certain African musical traditions.

References

Baily, J. (1985). Music structure and human movement. In P. Howell, I. Cross, & R. West (Eds.), Musical Structure and Cognition: New Readings in the Experimental Psychology of Music (237-258). London: Academic Press.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. European Journal of Neuroscience, 13, 400-404.

Cadoz, C. & Wanderley, M. M. (2000). Gesture – Music. In M. M. Wanderley & M. Battier (Eds.) Trends in Gestural Control of Music (71-93). Paris, Fr: IRCAM – Centre Pompidou.

Campbell, L., Chagnon, M. J. & Wanderley, M. M. On the use of Laban-Bartenieff techniques to describe ancillary gestures of clarinetists. Manuscript in preparation.

Chion, M. (1994). Audio-Vision: Sound on Screen (C. Gorbman, Ed. & Trans.). New York: Colombia University Press. (Original work published 1990)

Clarke, E. F., & Krumhansl, C. L. (1990). Perceiving musical time. Music Perception, 7, 213-252.

Cook, N. (2003). Embodying performance: Sound, sight and gesture in the improvisations of Jimi Hendrix. Keynote Address presented at the International Conference on Music and Gesture 2003, Norwich, UK.

Cooper, G. W. & Meyer, L. B. (1960). The Rhythmic Structure of Music. Chicago: University of Chicago Press.

Davidson, J. W. (1993) Visual perception of performance manner in the movements of solo musicians. Psychology of Music, 21, 103-113.

Davidson, J. W. (1994). Which areas of a pianist's body convey information about expressive intention to an audience? Journal of Human Movement Studies, 26, 279-301.

Davidson, J. W. (2001). The role of the body in the production and perception of solo vocal performance: A case study of Annie Lennox. Musicaie Scientiae, V (2), 235-256.

Davidson, J. & Correia, J. S. (2002). Body movement. In R. Parncutt & G. E. McPherson (Eds.), The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning (pp. 237-250). Oxford: Oxford University Press.

Delalande, F. (1988). La gestique de Gould: Élements pour une sémiologie du geste musical. In G. Guertin (Ed.), Glenn Gould Pluriel (85-111). Louise Courteau, editrice, inc.

Delalande, F. (1990). Le geste musical, du sensori-moteur au symbolique. In Proceedings of the Second Colloquio Internazionale di Psicologia della Musica, Ravello, Italy.

Deliège, I. & El Ahmade, A. (1990). Mechanisms of cue extraction in musical groupings: A study of perception on Sequenza VI for viola solo by Luciano Berio. Psychology of Music, 18, 18-44.

Dittrich, W. H., Troscianko, T., Lea, S. E. G., & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. Perception, 25, 727-738.

Fodor, J. A. & Bever, T. G. (1965). The psychological reality of linguistic segments. Journal of Verbal Learning & Verbal Behavior, 4, 414-420.

Fredrickson, W. E. (1995). A comparison of perceived musical tension and aesthetic response. Psychology of Music, 23, 81-87.

Fredrickson, W. E. (2000). Perception of tension in music: Musicians versus nonmusicians. Journal of Music Therapy, 37 (1), 40-50.

Frego, D. R. J. (1999). Effects of aural and visual conditions on response to perceived artistic tension in music and dance. Journal of Research in Music Education, 47 (1), 31-43.

Gabrielsson, A. (1999). Music performance. In Diana Deutsch (Ed.), The Psychology of Music (501-602), 2nd edition.

Gabrielsson, A. (2003). Music performance research at the millennium. Psychology of Music, 31 (3), 221-272.

Gregory, A. H. (1978). Perception of clicks in music. Perception & Psychophysics, 24, 171-174.

Krumhansl, C. L. (1996). A perceptual analysis of Mozart's Piano Sonata K. 282: Segmentation, tension, and musical ideas. Music Perception, 13 (3), 401-432.

Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. Canadian Journal of Experimental Psychology, 51 (4), 336-352.

Krumhansl, C. L. & Schenck, D. L. (1997). Can dance reflect the structural and

expressive qualities of music? A perceptual experiment on Balanchine's choreopgraphy of Mozart's Divertimento No. 15. Musicae Scientiae, 1 (Spring), 63-85.

Lerdahl, F. & Jackendoff, R. (1983). A generative theory of tonal music. Cambridge, MA: MIT Press.

Levelt, W. J. M. (1989). Speaking: From Intention to Articulation. Cambridge, Massachusetts: MIT Press.

Levitin, D. J., Nuzzo, R., Vines, B. W., & Ramsay, J. O. (2005). Introduction to functional data analysis. Manuscript submitted for publication.

Levitin, D. J. & Menon, V. (2003). Musical structure is processed in 'language' areas of the brain: A possible role for Brodmann area 47 in temporal coherence. NeuroImage, 20, 2142-2152.

McAdams, S., Vines, B. W., Vieillard, S., Smith, B. K. & Reynolds, R. (2004). Influences of large-scale form on continuous ratings in response to a contemporary piece in a live concert setting. Music Perception, 22 (2), 297-350.

McFarland, D. H. (2001). Respiratory Markers of Conversational Interaction. Journal of Speech, Language, and Hearing Research, 44 (1), 128-143.

McNeill, D. (1992). Hand and Mind: What Gestures Reveal About Thought. Chicago: University of Chicago Press.

McNeill, D. (Ed.) (2000). Language and Gesture. Cambridge: Cambridge University Press.

Modler, P. & Zannos, I. (1997). Emotional aspects of gesture recognition by a neural network, using dedicated input devices. In Proc. KANSEI - The Technology of Emotion Workshop, Genoa, Italy, 1997.

Orio, N. (1999). The timbre space of the classical guitar and its relationship with the plucking techniques. In Proc. of the 1999 International Computer Music Conference (391–394). San Francisco, CA: International Computer Music Association.

Planalp, S., DeFrancisco, V. L. & Rutherford. D. (1996). Varieties of cues to emotion in naturally occurring situations. Cognition and Emotion, 10 (2), 137-153.

Quek, F., McNeill, D., Bryll, R., Duncan, S., Ma, X.-F., Kirbas, C., et al. (2002). Multimodal human discourse: Gesture and speech. ACM Transactions on Computer-Human Interaction, 9 (3), 171-193.

Ramsay, J. O. & Silverman, B. W. (1997). Functional Data Analysis. New York: Springer-Verlag.

Repp, B. H. (1993). Music as motion: A synopsis of Alexander Truslit's (1938) Gestaltung und Bewegung inder Music (Shaping and Motion in Music). Psychology of Music, 21, 48-72.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. Trends in Neurosciences, 21 (5), 188-194.

Runeson, S. & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectations, gender, recognition, and deceptive intention. Journal of Experimental Psychology: General, 112, 585-615.

Sheets, M. (1966). The Phenomenology of Dance. Madison: University of Wisconsin Press.

Stern, D. N. (1999). Vitality contours: The temporal contour of feelings as a basic unit for constructing the infant's social experience. In P. Rochat (Ed.), Early Social Cognition: Understanding Others in the First Months of Life (67-80). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., & Levitin, D. J. (2005). Cross-Modal Interactions in the Perception of Musical Performance. Manuscript submitted for publication.

Vines, B. W., Levitin, D. J. & Nuzzo, R. L. (in press). Analyzing temporal dynamics in music: Differential calculus, physics, and functional data analysis techniques. *Music Perception*.

Vines, B. W., Wanderley, M. M., Krumhansl, C. L., Nuzzo, R. L. & Levitin, D. J. (2004). Performance gestures of musicians: What structural and emotional information do they convey? In A. Camurri & G. Volpe (Eds.), Gesture-Based Communication in Human-Computer Interaction (468-478). Berlin: Springer Verlag.

Wanderley, M. M. (1999). Non-obvious performer gestures in instrumental music. In A. Braffort, R. Gherbi, S. Gibet, J. Richardson, & D. Teil (Eds.), Gesture-Based Communication in Human-Computer Interaction (37-48). Springer-Verlag.

Wanderley, M. M. (2002). Quantitative analysis of non-obvious performer gestures. In I. Wachsmuth & T. Sowa (Eds.), Gesture and Sign Language in Human-Computer Interaction (241-253). Springer Verlag.

Wanderley, M. M., & Depalle, P. (2004). Gestural control of sound synthesis. Proceedings of the IEEE, 92 (4), 632-644.

Wanderley, M. M., Depalle, P., & Warusfel, O. (1999). Improving instrumental sound synthesis by modeling the effects of performer gesture. In Proceedings of the International Computer Music Conference (ICMC'99) (418-421).

Wanderley, M. M., Vines, B. W., Middleton, N., McKay, C., & Hatch, W. (in press). The musical significance of clarinettists' ancillary gestures: An exploration of the field. *Journal of New Music Research*.

Williamon, R. A. & Davidson, J. W. (2002). Exploring co-performer communication. Musicae Scientiae, VI (1), 1-17.