

One voice or many?: Compound melody and auditory streaming

Bryony Buck

Centre for Music Technology, University of Glasgow, U.K.
B.Buck@elec.gla.ac.uk – <http://cmt.gla.ac.uk/>

Graham Hair

Centre for Music Technology University of Glasgow, U.K.
graham@n-ism.org – <http://www.gla.ac.uk/music/ourstaff/grahamhair/>

Nick Bailey

Centre for Music Technology University of Glasgow, U.K.
nick@n-ism.org – <http://www.n-ism.org/People/nick.php>

In: C. Tsougras, R. Parncutt (Eds.)

Proceedings of the fourth Conference on Interdisciplinary Musicology (CIM08)
Thessaloniki, Greece, 2-6 July 2008, <http://web.auth.gr/cim08/>

Background in Auditory Perception.

Auditory perception, speech analysis and attention studies suggest limitations to the amount of auditory stimuli an individual may perceive at any given moment. The question arises, therefore, as to the implications of this for the music-theoretic tradition; that tonal context may lead us to interpret successive events as two or more auditory informational streams. Studies in music perception suggest that attention and segmental streaming may be directed by changes in pitch, timbre, rhythm, and volume of a note. To investigate the neurological relevance of multiple melodic voices and structure in a musical composition, we need to better understand the relationship between auditory streams and global auditory perception within pieces as diverse as Bach's movements for solo violin and Webern's Saxophone Quartet. The vast majority of research into musical listening uses pure tones or manipulated sound as stimuli, manipulating frequency and amplitude but not considering the effects on the *music* under investigation. Recent investigations show striking differences in cortical activation between listening to tones or simple scales and listening to actual music (Parsons 2005). This illustrates a real need to conduct research into musical listening and perception that uses actual, unadulterated musical stimuli. In doing so the current paper also strives to bridge the disciplinary divide evident within music research, recognizing the need to acknowledge the constraints of all disciplines involved (namely here music, psychology, and electronic engineering). This investigation is expected to illuminate the function of auditory processes and the limitations of the auditory system regarding musical listening. It moves beyond previous research by employing actual musical stimuli rather than focusing on pure tone or manipulated sound. In this respect auditory streaming and segmentation is considered regarding the musical listening experience as a whole, focusing on the differences between 'music as composed' and 'music as heard'.

Background in Music Analysis

'Compound Melody' is a ubiquitous feature of structure in tonal music, and music-theoretic explanations of passages which embody the concept cannot but be – at least implicitly – theories of listening, not just translations from musical sounds into words of what happens in the music. But it is by no means clear how analyses of musical objects using strategies derived from music-theory relate to theories of listening derived from strategies from perception.

Aims

We examine the capacity of auditory attention to assign non-consecutive events to the same auditory stream. Using cadentially segmented samples of actual musical pieces, we will investigate whether listeners perceive polyphonic melodies played by one instrument as either one 'voice' with two tones, or as two musical 'voices' (e.g. in phrases from J.S. Bach's Sonatas & Partitas for solo violin). In this way we wish to discern whether listeners perceive music in the way musical analysis denotes.

Main contribution

Participants with varying musical expertise and knowledge are presented with musical excerpts taken from J.S. Bach's Sonatas & Partitas for Solo Violin (BWV1001-1006). Musical stimuli have been segmented by cadencies, according to music-analytical theory. A two-alternative forced-choice (2AFC) task is designed to investigate whether participants perceive musical stimuli as comprising of one melodic voice with one or more melodic parts, or as multiple separate & distinct voices. Subjects are presented first with exemplars of simple and compound melodic musical structures. For each musical stimulus participants must make a judgement as to whether the presented excerpt contains 'one voice' or 'many voices'. Reaction time and decision response are recorded.

It is hypothesised that listeners do perceive musical structure such as compound melody in the same manner that music analysis denotes. The accuracy of this perception most likely correlates to an individual's level of musical knowledge and expertise, but is expected to be present even in non-musical control participants. Preliminary data supports this hypothesis. Implications and possible avenues for future research are discussed, with consideration to developing a new perceptual model of musical listening based upon music-theoretic analysis.

Implications

This research highlights the differences of segmentation, both horizontal and vertical, between tonal and atonal music. Perceptual effects are illustrated on a cognitive and neuronal level. Regardless of the experimental outcome, the results of this investigation will be informative both for composers and those researching auditory perception.

Musical Analyses: A disciplinary divide

Music as 'composed' versus music as 'heard' divides disciplines researching music and its effects. Attempts to reconcile these two aspects of music appear in discussions such as those of Eric Clarke (1986), Christopher Hasty (1977), Justin London (2004), John Sloboda (1997), and David Lewin (1993). As the investigation of musical research becomes ever more interdisciplinary, however, debates over technical terminology, experimental methods, and areas of importance seem to divide, rather than unite, the disciplines involved.

It has become apparent from such work that there is still a long way to go before research will understand exactly what the relationship between "music as composed" and "music as heard" is, or, to put it more generally, exactly what "listening to music" and "the perception of musical structure" actually involves.

Providing an answer to this problem is becoming increasingly important. From a Western musical perspective it has become commonplace, in recent decades, for criticism to claim that what happens in many contemporary compositions involves processes which "cannot be perceived". These claims imply that the status of processes which composers employ are no more than of autobiographical importance.

Music theory has a long history of endeavours to explain how music can be interpreted. These can be said to be, at least implicitly, theories of listening rather than theories of composition. The theory of "compound melody", to name only one of thousands of possible examples, is part of every theory of Western classical music devised since the middle ages (example figure 1) 'Compound Melody' theory suggests that non-successive notes in a melody must be connected in the listener's "mind's ear" in order that the music's structure be understood.

Recently, music theory has begun to seek the help of various branches of science to investigate some of its central problems. However, there are several factors obstructing such efforts, which are apt to produce misunderstandings on both sides. Music theory has its own language and concepts which are often difficult for non-musical scientists to negotiate. As a result, music theorists often complain of the naïveté of the way in which the majority of scientists approach musical tasks, and of their concepts of music more generally. Conversely, because music-theoretical concepts have not been often or widely tested empirically, nor tested using methodologies recognised by scientific communities, scientific researchers may complain that music theory holds no status as a form of empirical investigation.

Research into auditory perception of sound (as opposed to speech) has only attracted scientific over the last decade, particularly within the realms of neuroscience and cognitive perception. Computational and electronic investigation into music structure its components has interested audio engineers for only a very few decades. In contrast, the structural analysis and music theoretic models of today are grounded in over 200 years of practiced musical tradition.

Despite such a breadth of research, these investigations rarely overlap. A disciplinary divide exists preventing findings, models and theories from becoming accepted or accessible across disciplines and research a.

Inter-disciplinary discussion on the issue seems always to stumble over the insistence from both sides that only the methodologies entrenched in their own communities have any real validity. As a result of this, as researchers, we are no closer to understanding what music *is*.

Most notably, the current research begins with music as its foundation. Although a seemingly logical first step, this approach is counter to the majority of research in operation today, both perceptual and computational. In the main, psychological investigation into the effects and percepts of music utilises musical sounds as stimuli, such as alternating pure tones (ABA etc.). These are then altered and manipulated across factors such as frequency, pitch, timbre, meter and rhythm so as to investigate the individual effects of these elements on perception. This removes the fundamental core of 'what music is' from the thing being researched, leaving the findings relevant only to perceptions of sounds rather than of music.

Equally within computational & electronic research, in the majority of approaches, algorithms are devised to which music is then presented; little or no sophisticated music theory is performed *ante hoc* to inform algorithmic design, placing music itself last in the investigation process. Within our research team we believe that these research ideas, no-matter how novel or potentially revealing, cannot be fruitful without considering music, as it stands and exists, from the very beginning of research. Deriving technology, theories, and research methods from music-analytical processes, we hope to realise the piece of the puzzle that has proven so elusive throughout previous research.

'Compound Melody' is a ubiquitous structural Auditory streaming has been examined using

Figure 1. Exemplar of complete contrapuntal compound melody in Bach fugue analysis. The different colours represent different melodic voices running throughout the piece. The voices demonstrated here are not as ambiguous or as precise as the compound melody that exists within our musical stimuli. However, this serves extremely well to demonstrate what is meant by melodic voices within musical structure of a piece.

feature of tonal music, and music-theoretic explanations of passages which embody the concept cannot but be, at least implicitly, theories of listening: simply translations from musical sounds into words describing what happens within the music. It is by no means clear, however, how analyses of musical objects, using strategies derived from music-theory, relate to theories of listening derived from strategies of auditory perception.

This research examines compound melody and auditory streaming (also considered a process of perceptual organisation). In doing so the psychological elements of music perception may be revealed, making advances in theories of musical listening and computational search and not only make

match-programmes more plausible but also their application highly effective.

Auditory Streaming

The processes of streaming auditory information and of perceptual organisation strongly effect how we hear and understand music. Indeed, to create certain rhythms, harmonies and melodic patterns, composers will explicitly manipulate compound melody and other cues relevant for such processes (Huron, 1991).

Various theories have been proposed for perceptual recognition and processing of auditory streams as regards musical sound. However, little is known about the underlying processes involved, nor about what makes music so different from other auditory stimuli such as vocalised, environmental, or computer generated sound.

various methods (Sloboda, 1998) including behavioural science (Repp *et al.*, 2005) cognitive neuroscience (Cusack, 2005; Overath *et al.*, 2007; Brown *et al.*, 2006; Stewart *et al.*, 2006; Stainsby *et al.*, 2004; Peretz & Colethart, 2003; Zatorre *et al.*, 2002; Khalfa *et al.*, 2005; Watt & Quinn, 2008; Kuriki *et al.*, 2006) language studies and computer modelling (Cont *et al.*, 2007).

In addition, compound melody has been discussed in depth within musical theory & analysis (London, 2004; Lewin, 2003; Deleige & Sloboda, 1997), education (Beuvois & Medis, 1996; Bamberger, 2003; 1991) and music perception (Lewin, 1993; Huron, 2006; Zibowski, 2002; London, 2002).

Rhodri Cusack and colleagues (2005) utilised functional magnetic resonance imaging (fMRI) in order to identify an effect of perceptual organisation across the brain as a whole. They presented participants with an 'ambiguous auditory figure' that could be perceived sometimes as two streams, sometimes as one. Employing methods used for visually ambiguous stimuli, Cusack's auditory stimuli were compared across physically identical presentations so as to match the demands placed upon basic processing. The stimuli were designed in such a way that, across time, the percept would alternate randomly between the perceived states: one auditory stream or two (Cusack *et al.*, 2005; Parsons, *et al.*, 2005). They found that when two streams were perceived there was an increase in activity in the intraparietal sulcus (IPS) compared to when one stream was perceived. This involvement of the IPS in binding and organising auditory

information supports previous research into auditory streaming in both humans (London, 2002; Gregory, 1994; Hartmann, & Johnson, 1991) and macaques (Micheyl *et.al.*, 2005), showing a distinction in cognitive processing between the perception of single and multiple auditory streams.

When considered alongside music-theoretic analysis of compound melody, such research raises the question 'Do listeners actually *hear* multiple melodies that run within a musical piece in the same manner that music analysis denotes?'

Research into duet singing in fMRI (Parsons & Osherson, 2001) investigated complex music vs. simple song using human-human and human-computer duet conditions (as compared to solo singing). Results showed an intricate network of activation when duetting with the computer that differed dramatically in comparison when duetting with a human partner. They concluded that complex music activates inferior frontal regions differing from those areas activated by simple song (Parsons-*in prep*; Parsons & Osherson, 2001; Raichle *et al.*, 2001).

Further investigation using fMRI for processing of musical scales compared to a Bach concerto (Parsons-*in prep*) revealed significant differences in activation and, more interestingly, in deactivation during processing for the two conditions. This deactivation may be due the increase in concentration required for more complex and informative stimuli (Bach), following findings that cortical deactivation occurs during concentration (Raichle *et.al.*, 2001).

Similar activity is found when a person is passively listening to music and when generating music, although activation is stronger during the latter (Parsons-*in prep*). This pattern is common in both musicians and non-musicians when generating and listening to music for meter, tempo, rhythm and pitch. Musical tasks appear supported dynamically throughout the brain (Parsons-*in prep*; Parsons & Thut-*in prep*; Cusack, 2005) and appear task dependent.

Recent research by Parsons and colleagues clearly demonstrates a difference in cortical activity and, more interestingly, deactivation during processing for actual musical excerpts as opposed to repeated tones or simple scales (Parsons & Thut-*in prep*).

In functional neuroimaging studies non-musicians have been found to generate similar but asymmetrical activity in the left and right hemispheres for speech and singing processes respectively (Riecker *et.al.*, 2000; Jeffries *et.al.*, 2003).

Such results reflect concerns among the musical community that psychological

research, despite its claims to the contrary, is not specifically investigating the perception of '*music*' per se, but rather that musically generated.

The Parsons, Brown, and Dunbar research teams have shown differences between processing simple tones and musical scales compared to actual musical pieces. These findings support not only the current paper's use of musical excerpts taken from J.S.Bach's Sonatas & Partitas for solo violin (BWV1001-1006) as auditory stimuli, but also highlight the need for psychological, perceptual and computational research to use music in order to study it.

Music-theoretic Approach

Jean Bamberger highlights the paradoxical nature of music theory, particularly in her quotation of Louis Krasner¹ to a student, instructing her to "Forget about the notes and play the music"(1996: p35)

The student in this case needed the notes in order to play the music, and therein lays the problem of music theory, as Bamberger describes it. Context of understanding and learned descriptions inevitably carry through into our experience. Bamberger points out that as a person gets older they begin to lose the abstract nature of their perceptual experiences. Instead they generally report what were initially only learned descriptives of an experience as being perceptual (Bamberger, 2003; 1986). Learning informs experience. This highlights what Bamberger calls a "tacit ontological commitment" proposing we reconstruct our abstract perceptual experiences using the learned descriptives of that phenomenon.

The children in Jean Bamberger's studies discovered structural features and compositional methods similar to those described by music theory through their own abstract means. Based on these findings she has devised a music-learning program which quite literally turns "music theory on its ear" (Bamberger, 1996; 1991). By allowing students to utilise a top-down method of analysing music, Jean Bamberger provides evidence to suggest that we may indeed perceive structural features such as phrase ending and pitch contour that are prevalent within music theory without formal training.

Bridging the Divide

The current paper aims to follow Jean Bamberger in "turning musical theory on its ear" whilst at the same time turning psychological theories of music perception on its head.

¹ An eminent violinist who in later years taught at the New England Conservatory of Music until his death in 1994.

By presenting unmodified musical excerpts of J.S.Bach's Sonatas & Partitas for solo violin, stimuli will remain constant in musical source and timbre (so satisfying psychological experimental constraints). Excerpts are extracted according to music analytical phrasing, with each excerpt consisting of one phrase, divided between cadences.

The current study elected to use the Bach solo violin suites as they are commonly used for pedagogical reasons, partly because they exemplify highly formalised genres of composition and music structure. As they are written as solo pieces for the violin, there is only one source of sound and so timbre remains relatively constant throughout the stimuli. However, unlike some other instruments, the violin is capable of playing multiple notes at once.

Although these pieces exemplify compound melody in a fairly obvious way (within music theory) they are not entirely contrapuntal, making use of harmonic structuring to develop the voices within the main tune. As a result the compound melodies running throughout these selected pieces are often subtle and ambiguous, running in a fashion that may at first appear counter-intuitive. The quantity of information provided through the music is substantial for the musically experienced listener. Anyone familiar with music of this genre (18th century German Baroque) will recognise the community-folk-dance style structure coming through in many of the movements used.

The experiment follows a relatively simple design, not dissimilar to that of Michely et al., (2005). The task employs a 2-alternative forced-choice task (2AFC) where participants make judgments as to whether a given excerpt contains one or more melodic voices.

It is the informational content and the task demands that make this such a difficult and complex task. The task assumes a knowledge of not only musical theory and analysis, but an implicit awareness of compound melody in order to perform the task 'correctly'. In reality there is no correct or incorrect perceptual response, as the experiment is designed to ascertain whether or not musicians implicitly hear musical information in the way music-theory and compositional analysis proclaims.

The aim of the current research is to discover whether psychological behavioural methods can determine if musicians perceive compound melody within given musical excerpts. If results demonstrate a correlation between musical analysis of compound melody and perceived melodic voices, it will be possible to infer that music theory is not simply autobiographical explanation of compositional structure, but a theory of listening and perceptual organisation with regards to compound melody and auditory

streaming. Such findings would not only suggest a new (psychologically) and well founded (in musical theory) path for future research to take, but would also go a long way towards bridging the disciplinary divide currently limiting the progression of musical research.

Methodologies

Participants

Initial data was recorded using six volunteers with musical knowledge ranging from 'minimal training and no performance experience', to 'amateur musician and performer'. Participants were all postgraduate students at the University of Glasgow and all participated with informed consent.

Based on preliminary results, the main body of this research (currently underway) will use two test groups of musicians (highly experienced and amateur performers). In addition a control group of non-musicians will be tested in order to demonstrate that results are not simply an effect music-theoretic learning, but occur in individuals without knowledge of music-analysis. The scheduled participants for the experimental phase are as follows:

20 musicians are participating voluntarily in the experiment, having first given informed consent.

Of these 10 are professional musicians with approximately 24 years musical experience and music-theoretic learning. The remaining 10 participants are musical experts with 10-20years musical experience and a minimum of 4 years formal theoretical training.

Participants are aged between 20-65, with the mean age for musical professionals being (30-40)years, and for amateur musicians being 23years.

To account for the group condition, a control group of 10 non-musical participants will be recruited. Participants are matched to test-participants for handedness and age but have received no formal training within the last 10 years and none have achieved training or performance experience beyond amateur level (majority have received no musical experience whatsoever).

Fig.4. Examples of the first four cadential phrases extracted from Bach's corrente movement in the 1st partita for solo violin (phrases 1-4: P1-4). Each phrase ends at the closing tonic of the cadence. In phrase 4 an example of the melodic voices within the phrase is given with the base voice in blue and the melodic voice in red. According to music theory, in trials where only one melodic voice is present participants would be expected to report 'one', whereas in instances where two or more voices exist participants would report 'many'. However, often times these voices converge, cross or even disappear temporarily. It is the view within Schenkerian analysis that one must retain these base and melodic voices in the 'minds ear' in order to understand the musical content.

Stimuli & Apparatus

Stimuli comprised of approximately 126 excerpts from J.S.Bach's Sonatas & Partitas for solo violin (BWV 1001-1006). Excerpts were generated using music-analytical decompartmentalising of phrasing based on cadences. Segments run from the beginning of a phrase up to and including the first cadence (partial and formal) in that phrase. The second section would then run from the end of this cadence up to and including the next cadence in the piece (examples –figure two).

Six excerpts comprised the practice block, all selected for their clear, unambiguous structure of compound melody. The remaining 120 excerpts constituted the test phase of the experiment. All 120 excerpts were presented twice –once during the reaction time response task, and once during the accuracy task.

Musical stimuli duration is approximately 10-20seconds, depending on phrasing structure employed. Mean duration is ~13seconds. Stimuli were prepared using Audacity software. Original recordings came from 'Sonaten & Partiten [CD2.2]' (partitas) and 'Bach - Sonatas und Partitas fuer Violine soloCD1' (sonatas) with Jascha Heifetz as performer [BWV 1000-106]

Participants were presented with a 'Ready?' button, followed by 40 musical excerpts (each approximately 13seconds in duration with a 2000ms silence after each excerpt). A break was given after every 40th trial (approximately every 10minutes).

All excerpts have been identified as containing one or two or many melodic voices.

The experiment was run using a Java Applet specifically written to meet the demands of the task. Participants were placed in a soundproofed room designed for linguistic and auditory research.

There is an additional application within this applet enabling wider distribution and testing via the internet, which may be monitored and controlled for computer server, musical expertise, and age.

Data is recorded in the form of RT(nanoseconds), key selection (decision judgment), and accuracy through a Linux serving Java applet.

All participants were played the same array of musical excerpts, presented in a fully randomised order. Response choices were fully counter-balanced (CB) in order to rule out any effects of handedness.

Procedure

Participants are asked to listen to a series of short musical excerpts and to make a decision as to whether each piece comprises of one or more melodic voices. As all participants except those in the control group have musical training to a substantial level, all were aware of the term 'compound melody' and the compositional theory

behind this.² Participants perform a two-alternative forced-choice judgement (2AFC) using pre-assigned keys on the keyboard ('A' and 'L'). Participants are reminded of which key corresponds to which response through the visual presentation of two buttons; to the left and to the right of the visual interface graphic of the applet.

On-screen, participants are presented with buttons relating to the choices 'one voice' or 'many voices' corresponding to the key-press CB condition. The audio file is played through professional-quality headphones

while participants sat in a soundproofed room³.

During the first phase of the experiment participants are asked to respond as quickly and accurately as possible. The full array of audio files are then presented a second time (again in fully randomised order). However, in this second phase participants were told they may take as long as they wish and to be sure that they have answered correctly. In this phase only the control-group are expected to demonstrate a speed-accuracy trade off. The musical experts should show little difference in accuracy if making use of compound melody in their typical musical perceptions.

After the experiment, all participants are asked to demonstrate understanding of compound melody implicitly by asking them to sing the root progression of a single excerpt. Accuracy is assessed by Composer and Professor of Music, Graham Hair (composer in residence, Directa Nova & former Director of Music at University of

Glasgow – currently a research fellow within the Centre for Music Technology, University of Glasgow).

Measurements

Using a purpose-written java applet and dynamic html, we are able to record the user I.D., response key, response time (for both the depress and release of the response key), and the IP address and state of the computer environment on which participants perform the experiment. This has several benefits for investigation:

Primarily, the custom-designed program gives the researcher complete control over experimental conditions and constraints. Furthermore, as the experiment presents on a web-based browser applet, it is possible to invite musicians to participate from any computer connected to the internet.

The advantage here lies in the test group constraints and participant numbers that become possible when laboratory constraints are lifted. It is possible to empirically test musicians that meet the groups requirements that may otherwise be unavailable to perform the experiment. Experimental design is such that minimal loss occurs in the lifting of laboratory time constraints and environmental controls.

Data Analysis

Reaction time (RT) and accuracy (acc) data are recorded and stored alongside the key-press decision response using a Java applet and a PostgreSQL database. Participant responses are then compared to formal music analyses of compound melody for the same musical excerpts. A between groups ANOVA is to be calculated. In addition, the correlation percentage between participant response and music-analytical notation for the number of voices present will be assessed.

The authors intend to perform follow-up analyses using Bayesian modelling methods in order to generate further statistic power within results.

Results & Discussion

In order to investigate the effect of compound melody on auditory streaming listeners were presented with musical excerpts taken from J.S.Bach's Sonatas & Partitas for Solo violin. Excerpts lasted approximately 10-15seconds in duration. Listeners were instructed to indicate whether each piece comprised of one melodic voice or many (more than one) using a 2AFC key-press response. In the initial phase participants were asked to respond as quickly as possible. During the second test-phase participants were asked to respond as accurately as possible, without time constraints.

Preliminary Results

A preliminary experiment was run using six volunteers, all with some degree of musical background (half met the criteria for amateur professional musician, with the other half falling into the group of 'other musician' signifying some musical training and knowledge but to a low degree). In this pilot

study subjects were presented with 33 musical excerpts via a desktop PC computer and professional quality headphones. Judgement responses were recorded as per the previous methods section.

Due to the small sample size and the nature of the pilot participants, preliminary results are merely suggestive and hold no conclusive statistical power.

Data indicates that listeners have no trouble detecting the upper melodic voice, particularly when this is the main voice within the piece. Base voices appear less obvious and are more difficult to detect. However, participant responses correspond with musical analyses at levels that are above chance (67% of trials). Over the full experimental stimuli set of 126 trials this percentage is expected to increase due to a larger and more accurate data set. Although participant responses agreed above chance with optimal responses identified by musical analyses, it appears that the structure of the compound melody may sometimes seem counter-intuitive to the participant; running in a different pattern than the ear perceives although still existing and presenting the same effect.

Preliminary results indicate a trend suggesting that participants are in fact able to perceive the structure of compound melody, as denoted by musical analyses, within musical stimuli. Although this trend supports the authors' hypotheses, results did not reach significance. This is undoubtedly due to the small size of the data set obtained. Wider research across the sample population sizes suggested in the 'experimental procedure' section is expected to generate statistical significance within results.

Following the trend of results that is indicated in the preliminary data analyses, several inferences may be made:

On comparing musical expertise (amateur professional musician versus individuals with a low level of musical knowledge & experience), judgement responses seem to correlate extremely closely to the optimal responses identified by music-analytical theory. For stimuli comprising multiple melodic voices (response 'many'), non-expert listeners' responses were congruent on ~75% of trials, and amateur musicians on 24/26 trials (99.92% agreement). Where the music-theoretic response was 'one voice' participants with low musical expertise agreed on 66.63% of trials and amateur musicians agreed on all trials. In addition to suggesting that participants may indeed perceive compound melody similar to that specified by musical analysis, results also appear to reflect an effect of musical expertise upon perception.

Amateur musicians responded faster than those with low musical experience for stimuli containing compound melody (mean reaction times: amateur musicians=7090.1ms; other musicians=7737.7ms). Conversely, the group mean reaction times are reversed for stimuli comprising a simple melody (No compound melody present) (mean reaction times: amateur musicians =9617.2ms; other musicians=6935.8ms). These results provide evidence of an interaction, but due to the lack of significance in preliminary results, it was not possible to perform an analysis of variance to discern statistically where this interaction lies.

However, considering the trend in the results, the current authors would suggest the following theoretical explanation for results: It stands to reason that participants with greater knowledge of musical structure and compositional tools would take longer to respond when a sample demonstrates only one melodic voice. Participants are likely to continue listening as long as possible to ascertain, and be certain, that an additional yet subtle melodic voice does not diverge from the main melody later in the excerpt. Alternatively, participants with little to no knowledge of compositional structure and compound melody are likely to base their judgement on the auditory stream as it is presented, so not showing expectation behaviour in the form of prolonged reaction times. This relates to theories of musical listening and learning based upon expectation postulated by Huron (2006).

Therefore, the data gathered seems to support the hypotheses that listeners able to perceive compound melody within ambiguous musical stimuli. Furthermore, initial data implies an indirectly proportional effect of musical expertise on reaction time, whereby participants with greater musical knowledge appear to respond faster when compound melody is present in an excerpt, but far slower when the sample comprises only one melodic voice. This decrease in reaction time is almost certainly an artefact of the experimental design, as listeners were explicitly listening for 'one' or 'many' melodic voices.

Individual differences between the participants make presenting conclusive data impossible at this point. A full and complete data set is required in order to infer any theoretical findings. Initial results do, however, indicate that musical listening may indeed follow a similar vein to that denoted by music-analytical tradition. Based on preliminary findings, music-analysis may well be the starting block from which to leave in the search for a new model of musical listening and perception.

General Discussion

Although this experiment appears relatively simple in design, the actual task set to participants is rather complex. It assumes a knowledge of not only musical theory and analysis, but an implicit awareness of compound melody in order to perform the task 'correctly'. In reality there is no correct or incorrect perceptual response, as the experiment is designed to ascertain whether musicians implicitly hear musical information in the way music-theory and analysis assumes.

At present only initial pilot data has been gathered. Web-invitation to the applet will follow full lab-based experimentation with empirical controls. The web-based research will serve to confirm the external validity of empirical findings.

Initial data indicates support for the hypothesis that structural features denoted by musical analysis may be perceived by the listener when processing the auditory information. Participants agreed above chance with music-analytical theory in the recognition of multiple melodic voices across trials. Responses were notably faster for detection of simple melodic structure (one voice) as opposed to that of compound melody (multiple/many voices) for listeners with low levels of musical knowledge and experience. An effect of expertise was observed as listeners with greater musical knowledge and experience responded faster than inexperienced participants when detecting compound melody, but were remarkably slower to respond when making 'one voice' judgements. Due to the sample and data size it is not viable to make conclusive inferences here. It is noteworthy, however, that these trends relating to reaction times go some way to support David Huron's proposal of a substantial role of expectancy in musical and auditory processing (2006). The response times for those with less musical knowledge also follow on from Parsons and colleagues' research into differential processing for simple versus complex auditory stimuli. Moving on from their results showing substantial differences in both cortical activation and deactivation when comparing simple scales and tones to processing of a Bach concerto, the trend observed here suggests this may continue for complexity of compositional music processed.

The current investigation has successfully, insofar as results are preliminary, combined music-analytical theory with psychological experimental methods to examine auditory perception of musical structure. Initial data indicates that wider investigation is necessary, however, the fact that it is possible to observe such interesting trends from such preliminary data illustrates the adequacy of the methodology and experimental design used. It must be remembered, though, that

results are only suggestive, and by no means conclusive.

The implications of this research are discussed below, giving consideration to all possible outcomes following full experimental investigation. At the present time, however, it would appear that results certainly lean towards supporting the authors hypotheses that listeners may perceive music using similar structural organisation to musical analytical theory. Therefore, musical analysis may well prove to be an ideal starting point from which to move towards developing a model of musical listening and auditory perception that is applicable across disciplines, so moving us closer to discovering what music *is*.

Concluding Remarks

This paper demonstrates the implications of possible outcomes, giving consideration to future research.

Implications of Possible Outcomes

Given the outcome that musical experts do in fact hear compound melody in the same manner that music analysis denotes, several research implications arise:

Such findings would cast doubt on the assertion that compositional structure is little more than an autobiographical tool used by composers. By demonstrating that structural features such as compound melody may be perceived during passive listening, we lend support to the idea that music theory may also provide a foundation for a theory of musical listening. Further investigations are necessary into other features of music analysis. To strengthen these inferences, investigation into the underlying cortical processing of structural features of music would also be pertinent. EEG and MEG analysis of temporal activity across the brain would not only enhance the claims above, but also provide further investigation into identifying the function and location of the musical auditory cortex.

The implicit perception of compound melody would also inform current theories of listening. Research suggests that musical and speech phrase-production may have similar processes but differing semantics (Zatorre et al, 2002). The outcomes of the current paper would indicate whether passive listening may have processing pathways similar to those well established in speech-listening. By demonstrating this using actual musical stimuli, these results are pivotal in combining existing research into speech and sound with music-theoretic models of music structure.

Further investigation into these findings will lead us closer to identifying a temporal and

functional model of musical listening.

In the event that non-musical controls do not perceive compound melody, results will indicate that perception of music-theoretic structure is informed by learning and musical expertise. This outcome supports Jean Bamberger's 'tacit ontological problem'; that we begin to hear the descriptives that we have learned, rather than the percepts that we hear (1996; 2003). Future considerations in this instance revolve around musical learning and memory, as well as the role of context upon understanding auditory stimuli.

Alternatively, if non-musical controls do implicitly perceive the compound melody within the musical stimuli, but to a lesser degree of accuracy than the musical experts, it may be inferred that compound melody is a structural feature within Western music culture, and a feature of musical perception. Such results would provide support for the existence of cognitive specialisation in expertise (Parsons, in-prep) as well as justifying further investigation using musical experts.

Regardless of outcome, the current investigation informs upon previous research into musically generated sound and the perception of music. Findings of this research move us, as researchers, closer to understanding the cues and processes involved in musical listening.

Such understanding will enable more encompassing models of musical listening to be identified. The key contribution of the current investigation, however, is its role in bridging the disciplinary divide that has dominated within musical and auditory research for decades.

Acknowledgements:

This is an EPSRC funded research project within the University of Glasgow's Department of Electrical and Electronic Engineering. Thank you to all those of you who are involved in the learning and transferring of skills essential to this inter-disciplinary research.

References

- Bach, J.S. '*Sonatas & Partitas for solo violin*'. Opus: BWV 1001-1006 / RCA B000003FIT - Jascha Heifetz (violin).
- Bamberger, J.S. (2003) '*The Development of Intuitive Musical Understanding: A natural experiment*'. *Psychology of Music*. Vol. 31(1) 7-36
- Bamberger, J.S. (1991) '*The Mind Behind The Musical Ear: How children develop musical intelligence*'. Harvard University Press. ISBN: 0674576063
- Bamberger, J.S. (1996) '*Turning Music Theory on its Ear: Do we hear what we see; Do we see what we say?*' *International Journal of Computers for Mathematical*

Learning. Vol.1, 33-55

Bamberger, J.S. (1986) '*Cognitive issues in the development of musically gifted children*' in Sternberg, R.J. & Davidson, J.E. (eds) (1986) *Conceptions of Giftedness* Cambridge University Press. 388-413

Beuvois, M.W. & Medis, R. (1996) *Computer simulation of auditory stream segregation in alternating-tone sequences*. Jnl Acoustical Society of America. Vol. 99, 2270-2280

Brown, S., Martinez, M.J., & Parsons, L.M. (2006) *Music and language side by side in the brain: a PET study of the generation of melodies and sentences*. European Journal of Neuroscience, Vol. 23, pp 2791-2803.

Clarke, Eric F. (1989) '*Mind the Gap: Formal Structures and Psychological Processes in Music*'. Contemporary Music Review. Vol 3/1, 1 – 13.

Cont, A., Dubnov, S., & Assayag, G. (2007) 'Anticipatory Model of musical style imitation using collaborative and competitive reinforcement learning' in Butz et.al., (eds) AbiALS 2006, LNAI 4520, pp. 285-306

Cusack, R. (2005) *The Intraparietal Sulcus and Perceptual Organization*. Jnl Cognitive Neuroscience. Vol.17(4), 641-65

Deleige, I. & Sloboda, J. (1997) '*Perception and Cognition of Music*'. Psychology Press Ltd. UK. ISBN: 0863774520

Gregory, A.H. (1994) '*Timbre and Auditory Streaming*'. Music Perception. Vol. 12(2), 161-174

Hartmann, W.M. & Johnson, D. (1991) *Stream Segregation and Peripheral Channelling*. Music Perception. Vol 9(2), 155-184

Hasty, C.F. (1997) '*Meter as Rhythm*'. Oxford University Press. NY. ISBN: 0195100662

Huron, D. (2006) *Sweet Anticipation: Music and the psychology of expectation*. MIT Press. ISBN:026282783

Huron, D. (1991) *Tonal consonance versus tonal fusion in polyphonic sonorities*. Music Perception. Vol. 9, 135-154

Jeffries, K.J., Fritz, J.B. & Braun, A.R. (2003) *Words in melody: An H₂¹⁵O PET study of brain activation during singing and speaking*. Neuroreport , 14 , 749 – 754 .

Kuriki, S., Kanda, S. & Hirata, Y. (2006) *Effects of musical experience on different components of MEG responses elicited by sequential piano-tones and chords*. The Journal of Neuroscience. Vol.26(15), 4046-4053

Lewin D. (1993) '*Musical Form and Transformation: Four analytic essays*'. Oxford University Press. NY. ISBN: 9780195317121

London, J. (2004) '*Hearing In Time: Psychological aspects of musical meter*'. Oxford University Press. NY. ISBN: 01951608190

London, J. (2002) *Cognitive Constraints on Metric Systems: Some observations and hypotheses*. Music Perception. Vol. 19(4) 529-550

Micheyl, C., Tian, B., Carlyon, R.P. & Rauschecker, J.P. (2005) *Perceptual Organization of Tone Sequences in the Auditory Cortex of Awake Macaques*. Neuron. Vol.48. 139-148

Overath T., Cusack R., Kumar S., von Kriegstein K., Warren J.D., Grube M., Carlyon R.P., & Griffiths T.D. (2007) *An Information Theoretic characterisation of Auditory Encoding*. PloS Biology Vol. 5(11) 2723-2732

Parsons, L.M (in Prep) Discussion of research and collaborations currently underway. Presentation on '*New Findings on the Brain Basis of Musical Skills and Dancing*' at University of Glasgow, Psychology Department, 25th January, 2008

Parsons, L.M. & Thut, (in Prep) Discussion of research and collaborations currently underway. Presentation on '*New Findings on the Brain Basis of Musical Skills and Dancing*' at University of Glasgow, Psychology Department, 25th January, 2008

Parsons, L.M., Sergent, J., Hodges, D.A. & Fox, P.T. (2005) *Brain basis of piano performance*. Neuropsychologia , 43 , 199 – 215 .

Parsons, L.M. (2003) *Exploring the functional neuroanatomy of music performance, perception and comprehension*. In Peretz, I. & Zatorre, R.J., (Eds) , The Cognitive Neuroscience of Music . Oxford University Press , Oxford , pp . 247 – 268 .

Parsons, L.M. & Osherson, D. (2001) *New evidence for distinct right and left brain systems for deductive versus probabilistic reasoning* . Cereb. Cortex , 11 , 954 – 965 .

Peretz, I. & Coltheart, M. (2003) *Modularity of music processing*. Nature Neuroscience. Vol.7, 688-691

Sloboda, J.A. (1998) *Generative Processes in Music: Psychology of Performance Improvisation and Composition*. Clarendon Press , Oxford .

Stainsby, T.H., Moore, C/J. & Glasberg, B.R. (2004) *Auditory Streaming based on temporal structure in hearing-impaired listeners*. Hearing Research. Vol.192, 119-130

Stewart, L., von Kriegstein, K., Warren, J.D. & Griffiths, T.D. (2006) '*Music and the brain: disorders of musical listening*'. Brain. Vol. 129. 2533-2553

Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A. & Shulman, G.L. (2001) *A default mode of brain function*. Proceedings of the National Academy of Sciences of the United States of America. Vol. 98, 676-682

Repp, B.H., London, J. & Keller, P.E. (2005) *Production and Synchronization of Uneven Rhythms at Fast Tempi*. Music Perception. Vol23(1), 61-78

Riecker, A., Ackermann, H., Wildgruber, D., Dogil, G. & Grodd, W. (2000) *Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum*. Neuroreport , 11 , 1997 – 2000

Zatorre, R.J., Belin, P. & Penhune, V.B. (2002) '*Structure and function of auditory cortex: music and speech*'. Trends in Cognitive Sciences. Vol. 6(1), 37-46

Zibowski, L.M. (2002) *Conceptualizing Music: Cognitive Structure, Theory, and Analysis*. Oxford University Press: N.Y.

¹ There are many examples of demonstrating and investigating this. I would particularly like to mention the following:
Deutsch (1975) 'Two-channel listening to musical scales', *Journal of the Acoustical Society of America*, 57: 1156-1160
Dowling & Harwood (eds) (1986) 'Music Cognition', Orlando: Academic Press
Gregory (1994), 'Timbre and Auditory Streaming', *Music Perception*, 12(2): 161-174
Huron (1991) 'Tonal consonance versus tonal fusion in polyphonic sonorities', *Music Perception*, 9: 135-154
Janata et al, (2002) *Science*
Sandell (1995) 'Roles for Spectral Centroid and Other Factors in Determining “Blended” Instrument Pairings in Orchestration', *Music Perception*, 12(2): 209-246
Sloboda & Edgeworthy (1981) 'Attending to two melodies at once: The effect of key relatedness', *Psychology of Music*, 9(1): 39-43
Zeki (1999) 'Inner Vision', Oxford University Press: Oxford

² See 'Measurements' section.

³ To ensure this was the case participants were given an aural post-experiment test where they were asked to sing the root progression of a musical phrase. This allowed us to determine whether participants understood the analytical theory and ascertain whether they may have employed such strategies during the experiment without making neither task nor the underlying music-theoretic procedures explicit, so as not to lead results.