



City Research Online

City St George's, University of London

Citation: Wang, Y. (2020). Finding my voice: an interdisciplinary and multi-methodological investigation into the relationship between performers' speech and musical expression. (Unpublished Doctoral thesis, Guildhall School of Music and Drama)

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/25615/>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

Finding My Voice

**An Interdisciplinary and Multi-Methodological Investigation
into the Relationship between Performers' Speech and
Musical Expression**

Yundu Wang

Guildhall School of Music & Drama

Research Department

**Submitted in Partial Fulfilment of the Requirements of the degree
of Doctor of Music**

April 2020

I declare that, except where explicit attribution is made, the work presented in this thesis is my own. I grant powers of discretion to the School Librarian to allow the thesis to be copied in whole or in part without further reference to the author. This thesis is 69,056 words, excluding the appendix and bibliography.

Yundu Wang

April 2020

Contents

Abstract	7
List of Figures	9
List of Tables	10
Acknowledgements	11
Chapter One: Introduction	12
1.1 Foreword	12
1.2 Introduction	15
1.3 Context	17
In the Practice Room	17
East Asia and the Piano: Success and Stereotypes	21
Exploring Language	26
1.4 Research Process and Questions	28
What's Missing in Current Research?	29
Methodologies	31
Experimental Research Focus	33
1.5 Outline of Thesis	34
Chapter Two: Literature Review	35
2.1 Introduction	35
2.2 Literature on the Language and Music Relationship	36
Scholarship Prior to the 21 st Century	36
Musicology and Music Psychology	38
Scholarship of the 21 st Century: A Debate	42
Evolution	45
2.3 The Connection between Speech and Musical Expression in Performance	46
Musical Prosody	47
Musical Expression and Interpretation	50
Universal and Biological Principles in Music and Speech: Emotion?	52

2.4 Rhythm	55
Intensity	59
2.5 Comparable Acoustic Variables in Speech and Music	60
Prominence and Boundaries in Speech	61
Pitch in Speech	62
Prominence and Boundaries in Music	62
Pitch in Music	64
2.6 Further Clarification: Speech Rhythm, Musical Rhythm, Musical Prosody, and Nuance	66
2.7 Research Review	68
Research on Expressive Timing in Music	68
Research on Speech Rhythm	70
Cross-Domain, Cross-Cultural Research	74
Finding Evidence for Speech Influence in Musical Performance: A Critique of Current Studies	78
Chapter Three: Experiment	85
3.1 Introduction	85
3.2 Experimental Design	87
Background and Overview	87
Practitioner and Insider Researcher	93
3.3 Participants	98
3.4 Criteria for Selecting Materials	100
Music	100
Speech	103
3.5 Pilot Study	103
Materials	104
Experimental Procedure	105
Music	105
Speech	107
Instrumentation	108
Annotation	108

Discussion and Variables	110
3.6 Main Experiment	112
Variables	113
Speech	113
Music	114
Possibly Confounding Variables Not Controlled For	114
3.7 Ethics and Limitations	114
3.8 Analysis and Results	118
Results and Statistical Analysis	120
nPVI_V and VarcoV in speech (English and Mandarin)	121
VarcoC and %V	124
Speech Rate and Overall Tempo	126
Correlations Between Rhythmic Metrics of Music and Speech	127
Summary of Statistical Findings	132
3.9 Post-experiment Auditory Impression Study	134
3.10 Discussion and Conclusion	137
Chapter Four: Autoethnography	146
4.1 Summer, 2018	146
4.2 Introduction	147
Dialogue 1: Autoethnographic Method	151
4.3 Practice and Musical Experiences: France and Belgium	156
Dialogue 2: Chamber Music	164
4.4 Speech in Practice: Germany, Autumn, 2018	168
Dialogue 3: Agency and Authority as a Performer	180
Chapter Five: East Asian Musicians in Western Classical Music	187
5.1 Introduction	187
5.2 A Universal Language?	188
5.3 Racial Marking and Stereotyping	192
5.4 The Model Minority	195
History of Western Classical Music in East Asia	196
East Asian Parents	204

My Story	204
East Asian Parenting	208
Immigration	215
5.5 Authenticity in Western Classical Music	220
Voices: Interviews with East Asian Classical Musicians	225
Categories of the Universalist view	225
Categories of the Particularist view	228
Chapter Six: Conclusion	233
6.1 Summary	233
6.2 Embodied Awareness	235
6.3 Revising and Reforming	237
6.4 Revitalising Connections	238
6.5 An Encore: Schumann Piano Quartet, <i>Andante cantabile</i>	241
Appendix A: Participant Consent Form	244
Appendix B: Participant Information Sheet	245
Appendix C: Questionnaires	247
Language Ability Questionnaire Part 1	247
Language Ability Questionnaire Part 2 English	249
Language Ability Questionnaire Part 2 Mandarin Chinese	251
Questionnaire for Bilingual Background- English/Standard Mandarin Chinese	253
Appendix D: Participant Profiles	256
Appendix E: Mandarin Chinese Simple Story	261
Appendix F: Timing Curves	262
Bibliography	268

Abstract

This thesis presents an interdisciplinary and multi-methodological investigation into the relationship between speech and musical performance, precipitating a personal exploration of the complexity of individual musical expression and its ties with the identity politics of Western classical music.

This research evolved in two stages, beginning with an interdisciplinary analysis of the relationship between speech and music through literature review and quantitative empirical experimentation. Drawing on empirical methodology from existing interdisciplinary research, I conducted a comparative study of the speech and musical performance characteristics of classical pianists with different native languages. Results seemed to suggest that native language may influence the musical expression of performance to a certain extent. But they also suggested that timing variability as an element of expression is highly individual, may affect both speech and music, and may not necessarily be expressed equally in both. This study also afforded an opportunity to clarify and disaggregate different systems and parameters involved in individuals' expression, both in speech production and musical performance, one that may be useful for further research in the area.

The second stage utilised a qualitative research method to reflect on the results of the experiment, as well as on my own musical practice and personal experiences: autoethnography. Departing from a review of ethnographic research on East Asian musicians in the Western classical music industry, I discovered that my personal experiences of straddling two different cultures have impinged on—and arguably left under-developed—my sense of identity and 'musical voice'. Finally, I argue that my own

understanding of musical expression has been shaped by established discourses that may no longer align—if they ever really did—with the current multicultural reality of ‘Western’ classical music.

This thesis contributes to original research in the growing field of artistic and practice-based research in music performance, as well as in those of phonetic science and cultural studies. At the same time, my work has been a personal journey towards a deeper understanding of the complexity and tensions between expressivity, authority, and identity within a world of mixed and overlapping cultural parameters. I hope that this study may be illuminating to those embarking on similar journeys, as well as to those responsible for preparing the way for the very many still to be undertaken.

List of Figures

Figure 1: First Line of Alban Berg's Piano Sonata, op. 1	18
Figure 2: 'Rosemary' by Frank Bridge (mm 1–8)	102
Figure 3: 'A Serious Case' (Rose, n.d.)	103
Figure 4: 'Yu Diao' by Chen Yi, mm 14–18	104
Figure 5: Textgrid, waveform, and spectrogram of Mandarin 1, sentence 1	109
Figure 6: Spectrogram of expressive performance by Mandarin 1, with quaver beats (vertical white lines) and quaver beat timing curve (horizontal white curve)	110
Figure 7: PMCC of L1 and L2 English VarcoV and Varco-m	129
Figure 8: PMCC of VarcoV (L1 English and L1 Mandarin) and Varco-m	130
Figure 9: PMCC of English (L1 and L2) VarcoC and Varco-m	131

List of Tables

Table 1: mean nPVI_V and VarcoV values for L1 and L2 English speech	121
Table 2: mean nPVI_V and VarcoV values for L1/L2 English (columns 2 and 3) and L1 Mandarin speech (columns 4 and 5)	122
Table 3: mean VarcoC and %V of L1 and L2 English, and L1 Mandarin speech	124
Table 4: mean rate of English and Mandarin read speech (syllables per second) and overall tempo (quavers per minute)	126
Table 5: mean nPVI and Varco scores of quaver beat durations	127
Table 6: mean %V, nPVI_V, VarcoV, and VarcoC of L1 English, L2 English, and L1 Mandarin speech; mean Varco and nPVI of music	128
Table 7: Statistical difference between language groups	132
Table 8: Intraspeaker correlation between speech and musical performance	133
Table 9: scores of auditory evaluation and musical rhythm measures	135

Acknowledgements

I am deeply grateful for the guidance and encouragement of my supervisors, Cormac Newark and Elinor Payne; through their wisdom and incredible wealth of knowledge, I have been challenged to achieve further than I ever imagined was possible. I am also immensely thankful for the guidance of Ronan O'Hora; his constant reminders to trust my instincts and search for what I truly desire in music have given me the strength to embark upon this journey. I would also like to thank Martin Roscoe, David Dolan, Charles Owen, and Vivian Weilerstein for their musical wisdom and belief in me, even at a time when I did not truly believe in myself.

I am further thankful for my parents; their deep love and unwavering support have carried me through several moments of extreme difficulty, and I am in awe of their strength and integrity. My brother has been an inspiration from the moment he was born, and although I am the older sister, I will never stop looking up to him. I am also grateful for the wonderful friends that I have made, both in music and in life. Among them: En Yuan and her deep intelligence, sensitivity, and empathy constantly remind me of what I truly value in life; Guang and her strength know no bounds, and the same goes for her caring nature.

Finally, I would like to thank Krzysztof for standing by my side and supporting me throughout this journey. His ability to withstand my periodic crises is remarkable, and I am truly grateful for his constant encouragement to learn more, to develop further, and to reach without fear towards my goals.

Chapter One: Introduction

1.1 Foreword

It is difficult to say when my fascination with the relationship between music and language began. As a child, I was encouraged to think of musical performance as a form of communication; by telling a story or expressing an emotion through playing 'classical' piano music, I felt able to communicate to my audiences without using words. This conception of music-making made me aware, early on, of certain similarities between speech and musical performance (for example, the end of a phrase of music and a phrase of speech can both be signalled by slowing down and getting softer). It also led me to consider my own speech (both verbal and musical), and the possibility that language was inherent in the music I happened to be playing. Given that the canon of classical piano music is mainly European, I eventually found myself faced with a predicament: if these pieces of Western art music were composed by people who spoke vastly different languages to my own, how could I, who am not European and not entirely Western either, effectively communicate this 'language-content' in my playing? Did I sound like a foreigner, with an accent that immediately revealed to others that I was different? Or, does the frequently heard phrase 'music is a universal language' express something more than feel-good cultural liberalism?

Before diving into these questions, a bit of background: I was born in Harbin, China, and immigrated together with my parents to the United States when I was five years old. Soon after my arrival, I began Kindergarten, where I had my first experience of feeling 'different'. Although I was at the age when many children's language skills weren't yet fully developed, I still vividly remember the frustration of being unable to communicate with the other children; the only English word I knew was 'No!'. For the next few years, I felt strange

and out of place while my English skills developed. Around the same time, I started learning the piano.

Although I was living in America, my family environment was very much Chinese. My mother's English was basic at best and my father had a strong accent in his speech, often with incorrect grammar. Mandarin was my primary language aside from being at school or with friends, and it was spoken loudly and boisterously. Even in the music I listened to, there was more Chinese than English lyrics: my mother's hobby and passion was singing Chinese pop songs of the 70s and 80s. There was always a karaoke machine with microphones attached to the television in every house that we moved to. Mandarin songs blasted throughout my childhood and adolescence nearly every day, as my mother was intent on learning and perfecting her beloved tunes.

When I turned fifteen, I became home-schooled so that I could have more time to focus on music. My parents found an online schooling programme, and for nearly three years, I was studying, practising, and living at home with little interaction with the outside world beyond the internet. Eventually, I re-entered the state school system for my final year, and subsequently at eighteen, I left my home in central Massachusetts to study at a renowned music conservatoire in Boston. Once there, my exposure to people other than my Chinese-speaking family increased drastically, and I flourished under the influence of inspiring teachers, talented peers, and challenging musical experiences. Interestingly, a significant portion of the student demographic at the institution were either Asian American or had recently come from East Asia. Unlike myself, these students had either been born in the U.S. or had grown up in their native countries and came to the U.S. for higher education. As I had been born in China yet raised in the U.S., I still felt slightly different from both these sets of peers.

Of course, there were students who had similar bi-cultural upbringings.

Unfortunately, I never managed to form close relationships with them, to the extent that I could discuss my feelings of difference. Regarding language, I also felt isolated because I had not met anyone who was as fluent in both English and Mandarin speech as myself. Those who emigrated when they were teenagers had an accent in their English speech, and others had either lost their Mandarin fluency or had not been raised to speak the language at all. These, and other subtle differences, made it difficult for me to identify with the East Asian demographic at my conservatoire. My friendship circles consisted of the two extremes: there were the Americans (mostly Caucasian) on the one hand, and the Chinese (whose home addresses were still in China) on the other. Because of this polarity of cultures, I continued to feel like the odd one out in both groups, and lacked a sense of belonging, as well as mutual understanding in terms of identity. My American friends (and sometimes even I) would jokingly call me a 'fake' Asian, while in the Chinese group, I was known as 'the American'; I straddled two perspectives, while remaining undefined. At the time, I was content with this lack of self-identity, and never thought about its potential impact upon my music-making. Almost inevitably, this changed when I left the States and moved to the United Kingdom for postgraduate studies.

In London, I was made to realise that, irrespective of my own lack of cultural identity—I could confidently identify only as 'musician'—I was being identified by those around me. For some (usually after hearing me speak), I was American, and for others who only had an impressionistic view, I was no longer just 'Asian' but, according to the local terminology, 'East Asian', 'Far Eastern', or even 'oriental'. It was surprising to me how definite my labels became once I left the country that I grew up in. I was also suddenly exposed to many more

cultures and languages than I had ever experienced before. My view of the world expanded beyond my expectations, and with that expansion, I found myself returning to the questions I had posed during my youth. Do I sound 'foreign' when I play Western classical music? Do my performances have as much expressive intelligibility (and therefore value) as those of someone playing music of their own culture, broadly defined? And just what, exactly, determines such value? Do I sound 'American' or 'Chinese' when I play, and if so how? Does any of this even matter?

1.2 Introduction

This thesis presents an investigation into questions that once seemed to me like youthful musings but have since developed into far-reaching concerns, ranging from the concept of expression as it applies to Western classical music to identity politics. My research journey began with an almost obsessive desire to understand the relationship between speech and music from a scientific standpoint, hence my delving into phonetic science, with brief references to psychology and neurobiology along the way. My initial goal was to prove that my doubts about my own musical expression had some kind of grounding in empirical reality; that I (and presumably other musicians who are native speakers of Mandarin) instinctively perceive, interpret, and execute this essentially European music differently from those who speak European languages. I reasoned that by exposing the existence of systematic musical difference based on language background I could confront and potentially discover ways of overcoming this disparity; I could finally become an authoritative interpreter of the canon. What I achieved, instead, was the discovery of more fundamental and personal concerns relating to the complexity of individual musical expression on the one hand and the identity politics of Western classical music on the other.

My research has, therefore, been a process of two stages, beginning with an interdisciplinary apprehension of the relationship between speech and music through literature review and empirical experimentation. This first stage was an attempt at answering my research questions by quantitative means; specifically, I investigated the performative elements of musical expression that may parallel prosodic elements in speech production. Drawing on empirical methodology from existing interdisciplinary research on the language-and-music relationship, I extracted and analysed certain acoustical elements involved in both speech and expressive musical performance. By taking a comparative approach, I effectively conducted a cross-linguistic and cross-domain empirical study on classical pianists with differing native languages. It was during the interpretation of my experimental results that I realised that this attempt at an ‘objective’ investigation raised more questions than provided answers, and that I needed to involve a more personal—and at times disconcerting—probe into my own musical practice and understanding of musical expression. This realisation led to the second stage, which involved utilising the qualitative method of autoethnography to reflect on the results of my experiment and on my musical practice and experiences. During this stage, I also conducted a review of ethnographic research on East Asian musicians in the Western classical music industry; to my surprise, I found myself resonating with, and being represented by, the work I found. I realised that my personal experiences of straddling different cultures have led to an under-developed sense of self-identity and of my ‘musical voice’. In conclusion, then, I argue that my own understanding of musical expression has been shaped by established discourses that may no longer align—if they ever really did—with the current multicultural reality of ‘Western’ classical music.

This thesis contributes to original research in the growing fields of artistic and practice-based research, along with music performance studies and phonetic science, with a nod towards cultural studies. My work has been a personal journey towards a deeper understanding of the complexity and tensions between expressivity, authority, and identity within a world of mixed and overlapping cultural parameters. I hope that my words may be illuminating to those embarking on similar journeys, as well as to those responsible for preparing the way for the very many similar journeys still to be undertaken.

In this chapter, I elucidate the context behind my research topic and questions, as well as provide a brief overview of current research. I then introduce my quantitative and qualitative methodologies and provide a somewhat more detailed outline of the entire thesis. Given my use of an autoethnographic process, one derived from reflection on my research as it proceeded, it seemed appropriate to interweave, through this and other chapters, narratives that provide further context for my personal reflections.

1.3 Context

In the Practice Room

Although my curiosity towards the relationship between speech and music extended as far back as my childhood, I never imagined that it would eventually lead to doctoral research. I suppose the change from mild interest to a desire for deeper understanding began when I was recommended, by a fellow music student in Boston, to watch some video lectures given by Leonard Bernstein at Harvard (Bernstein, 1973: the Norton Lectures). Yet, while Bernstein's words on the parallels between music and linguistic syntax and semantics were

fascinating and—in a way—validating to my own private musings on music-making, they were not the first steps of my arduous journey into research. The initial push actually came from a particularly frustrating practice session, where my confidence in my own ability to make music had reached a low point and I was desperately searching for something that could ‘fix’ my playing and pull me away from certain familiar and morose thoughts, related to my ‘lack of talent’ and ‘lack of the “right” kind of expressivity’. During this practice session, I was preparing for a performance of Alban Berg’s piano sonata, opus 1, and was dissatisfied with my playing of the opening phrase. I believed that the first few bars held a significant emotional charge, as well as the basic motivic material of the entire piece, and I wanted to do them justice. In particular, I wanted to portray a sense of vocal utterance which is suggested by the detailed use of phrase and dynamic markings, as well as the breath mark at the end of the phrase (see Figure 1). I did not like what I was doing and, frustratingly, did not know how to improve it.



Figure 1: First Line of Alban Berg's Piano Sonata, op. 1

My dissatisfaction led me to the suspicion that I was hearing the phrase incorrectly in my head. I based this on the concept that, in order to play expressively, one must first be able to hear the expressive music, internally; following Kodály, this is known as ‘inner

hearing'.¹ This led me to believe that, if what I was playing did not sound 'right', I must have been singing it 'wrongly'. As I sang the notes again in my head, I focused my judgement on those 'inner' sounds and noticed something: I seemed to be imposing speech-like sounds to every note. In other words, I was imagining each note to have a consonant-like onset and a vowel nucleus with no coda (similar to 'la' or 'da'). I also realised that each note was imagined as a monosyllabic word, regardless of the length of the note or whether some notes were grouped within a single motif in the score (see the first three notes in Figure 1).² I was also more focused on the beginnings of each note, shaping them into nuanced sounds that fit together to form the larger contours of the phrases written in the score. I concluded from these observations that I was singing the notes of Berg in Mandarin Chinese; my inner hearing was accustomed to the lexical tones of Chinese speech, and certainly not to the lexical stress of German, the native language of the composer. While these observations were impressionistic, they prompted further considerations, including whether I could adjust the 'language' of my inner hearing to better suit the music. The most plausible adjustment for this particular instance was to superimpose speech-like sounds of another language that I was fluent in and that seemed similar to German: English. My strong (but again impressionistic) sense was that English sounded similar to German, or at least more similar than to Mandarin Chinese.

¹ This is a fundamental concept of the Kodály Method, a pedagogical system that was inspired by the educational philosophy of Zoltán Kodály, that has become commonplace in piano pedagogy. For an introduction to the method, see Dobszay (1972).

² The first three notes form a basic melodic and rhythmic motif; see Adorno (1991) on the music of Alban Berg.

Returning to the practice room, I ‘sang’ the first phrase of the sonata in my head with English-like speech before playing it again on the piano; instantly, I felt a change in both the way the phrase sounded to my own ears, as well as my physical approach on the instrument. The phrase sounded more continuous and fluid, while having more variety of timing *rubato* and subtle changes in dynamics. Also, the durational difference between longer and shorter notes was more apparent. Physically, I felt less strain in my hands and arms due to the more frequent release of tension during certain notes. My wrists became more active and supple in motion, whereas previously they had been strained and, frequently, locked in certain positions during performance. I was reminded of the concept that the wrist, in piano-playing, is commonly associated with vocalic, ‘singing’ phrasing; fluid motion of the wrist creates variety in colour and shade, as well as allows for lift in certain moments, creating an effect of ‘breathing’.³

Encouraged by this change in my performance, I sought to investigate the relationship between speech and music, as well as the possibility that one’s native speech influences one’s musical expression. The positive transformation of my playing after considering the ‘language’ of my inner hearing prompted me to conduct further research. I became curious whether the musical expression of pianists other than myself were influenced by their language backgrounds. If so, I reasoned, this could broaden debates within the Western classical music community about self-expression and individuality, not to mention authenticity and interpretative authority. Furthermore, I went on to conclude, the possibility of musical expression being related to speech poses questions about the

³ Fryderyk Chopin believed that the key to ‘breathing’ in piano playing comes from the wrist: ‘the action of the wrist is analogous to taking breath in singing’; see Eigeldinger (1986: 113).

music, or rather the compositions, themselves. Did composers write music in their own native 'language', and what does that mean for the performers who interpret their work? Does a performer, whose native language is the same as the composer's, have a head-start over other performers? Is it necessary to obtain fluency in different languages as one studies different repertoire? As I progressed through my investigation, I came across the work of others with similar interest in these questions, not only within the field of music (performers and otherwise), but also speech sciences, psychology, and neuroscience. I also began to look inward at my own language background, wondering if it was possible that my music-making was conditioned by Mandarin Chinese. This extended, by logical implication, to thoughts about culture and race. Thus was my journey mapped out into territory that, prior to conducting research, I never thought I would explore.

East Asia and the Piano: Success and Stereotypes

As my research progressed, it became clear that my work was located within a wider framework of cultural, social, and political significance. To begin with, to conduct research on classical pianists with varying linguistic backgrounds and cultural origins is to address issues that are embedded within the history of the piano and its popularity in East Asia. The modern piano is a significant cultural influence around the world, particularly in East Asia. Japan was the first to 'lionise' the piano as a symbol of cultural modernity during the pre-war era, followed by South Korea, China, and Taiwan (Tokita, 2010). There are currently around 40 million children learning the instrument in China. China has become the world's largest producer and consumer of the piano, contributing 79.6% of the global piano output by 2013 (ResearchMoz, 2013; ResearchInChina, 2014).

Today, East Asian pianists strive for, and succeed in gaining, praise and recognition on both the international competition and concert scenes. At the XVII International Fryderyk Chopin Piano Competition in Warsaw, Poland, four out of the six finalists were East Asian, with the first prize awarded to Seong-Jin Cho of South Korea. Chinese pianists Lang Lang and Yuja Wang are some of the most successful classical musicians in the world, appealing to those within the international community of musicians as well as to the general public. Such acclaim in a relatively short period of time (the current elevated status and popularity of the piano, and of Western classical music as a prestigious pursuit throughout East Asia really began only after World War II; see Wang, 2015) has trained the spotlight on Asian pianists, resulting in the scrutiny of factors contributing to Asian success. These factors seem to relate to conventional ideas about what is required of the classical performer to master the modern piano. It is widely understood that the piano places enormous demands upon the performer: to learn a significant amount of complex repertoire, to overcome exceptional muscular challenges and achieve considerable virtuosity, to have commanding physical presence, to be faithful to the composed work, yet authoritative and distinctive in performance. These factors contribute to certain prejudices and expectations, specific to Asian musicians, that have already been set by society outside of classical music: the notion that Asians are studious, obedient, and academically successful, and the notoriety of their rigorous and goal-oriented parenting, have long since trickled into the world of Western classical music and created conventional impressions of Asian musicians as being ‘technically exceptional’ and ‘hard-working’. At the far end of all this lie certain powerful stereotypes and narratives that are potentially damaging to Asian musicians. Across the classical music world (which these days means the whole world), epithets such as ‘technically advanced’ and ‘hard-working’ have often come to mean ‘lacking in musical

expressivity' and 'uninspired'. This vocabulary, needless to say, influences how the world perceives the artistic identity of Asian musicians. Clearly, the development of one's own artistic integrity, as well as the qualities of creativity, freedom of expression, and autonomy, may be heavily encumbered by such social narratives.

I must clarify here that these widespread sociocultural stereotypes represent a significant topic in various branches of social studies, one that this thesis cannot hope to cover adequately. I have not undertaken any systematic ethnographic research on this aspect of my project. Although I reference the work of scholars who have conducted in-depth investigations in these fields, in Chapter Five, I do so on the basis of my own personal journey. Equally, I do not expect my own particular reflections to be indicative of those of all East Asian classical musicians. However, if some of my experiences seem familiar to those reading, I hope my work will be useful in facing some of the difficulties and complexities I have encountered.

As a Chinese-born classical pianist who was raised and educated in the West, I experienced pressures both to conform to and disprove certain stereotypes. The expectation of East Asian pianists to be technically exceptional is one of them. This may have to do with the specifics of training in the East compared to the West. In Chinese piano pedagogy, for example, there seems to be a significant emphasis on the mechanics of playing, particularly regarding the development of a sound technical foundation. As highlighted in Xu (2018), renowned teachers in China such as Guangren Zhou⁴ focus on training students with exercises and drills, as well as specific stretching exercises. During the

⁴ Zhou had taught Yuja Wang, as well as Zhaoyi Dan, who is another highly esteemed pedagogue in China and the teacher of the XIV International Fryderyk Chopin Piano Competition winner, Yundi Li.

first year of Undergraduate piano study at the Central Conservatory of Music in Beijing, students undergo a rigorous training plan of two-hour lessons, in which students are asked to perform technical drills for up to an hour. Within the lessons, five-finger exercises, as well as scales, double notes, arpeggios, octaves, and chordal exercises are used as training. Particularly for students with small hands, stretching exercises involving the high lifting of fingers during the playing of chords are used. This kind of training, which, according to Xu, is widespread among piano teachers in China, has created reverberations throughout the global classical music community. The successes of young Chinese students on the international competition platform, taken together with statements such as ‘Brilliant technique and great power are what’s most admired in China’ by esteemed Chinese musicians like Cai Jindong, have influenced the Western perception of Chinese classical music students to be ‘trained mostly in technique’ (see Johnson, 2010). My own experiences within Western music conservatoires suggest that there are no such lessons dedicated to technical proficiency at such a stage; students work with teachers individually and technique seems to be addressed based on need, rather than on a set curriculum.

The following anecdotes, one regarding my own musical training and background and the other of a public masterclass, showcase the effect that stereotypes have had upon my own musical development and experience:

My own musical training did not include rigorous technical plans; as a child, my piano lessons were with a private teacher and consisted of learning pieces, not drills. Thus, I spent the majority of my time ‘perfecting’ the performance of pieces, intended for recitals and competitions. What I lacked in technical proficiency, I compensated for with musical instincts and emotional involvement during performance. This, however, only carried me so far.

During my studies in Boston, I was confronted with the imbalance in my pianistic training. In one particular lesson, my piano teacher declared that, while my musical abilities were at a superior level, my technical level was lower and needed to be focused and improved upon. This teacher continued to discuss the abilities of certain students who did have superior technique (all of whom were trained in China) and determined that I needed more work with studies and exercises. At the moment of hearing these comments, it seemed to me that the Asian stereotype of 'technical' was something I had to live up to. I began to devote myself to improving my technique, focusing on the physical aspects of playing the piano in order to obtain 'Asian' technical brilliance. In retrospect, perhaps my teacher had not intended such a separation between technique and musicality. However, at that period in my life, I was lacking in experience, confidence, and knowledge; I was sensitive and impressionable, working within a highly competitive environment, and eager to overcome any weaknesses to improve.

A few years passed, and I had resituated myself in London. It was there that I encountered a different label, one that forced me not only to re-examine my own musical development, but seemed to add to my confusion and self-doubt as a performer. I played Mozart's Sonata for Two Pianos with my duo partner (who is Japanese) in a masterclass for a highly-esteemed musician. The piece was new to us, and we hadn't yet settled on a convincing interpretation. The musician, after deeming our performance technically impressive but not characterful enough, immediately spoke about how difficult it is for Asian musicians to play with humour or extremes of emotion because, in Asian culture, such expressivity is not encouraged. I suddenly found myself facing the Asian stereotype of being perceived as 'closed', 'timid' and 'expressionless', if not (in the classic racist cliché) 'inscrutable'. It was a prejudice that I thought I was immune from, since I was raised in the

U.S., and no one had, up to that point, commented upon my musical performance by use of cultural generalisations (interestingly, this musician was also American). I was deeply puzzled as to why the musician chose to associate our lack of characterful playing with a broad cultural stereotype, rather than focus on the elements in the music that could be improved upon; I also felt that, for the audience in the hall at that moment, stereotypes were being reinforced, regardless of the validity of the statement. Whether or not I actually had the ability to play with humour and was simply having an 'off day' was potentially made irrelevant by the fact that I was—or at least looked—Asian.

Exploring Language

Such experiences, in combination with my frustrations during practice, as well as an increased exposure to the diversity of the world's languages since living in London, compelled me to investigate language, particularly speech, in more depth. I was led by the thought that there might be something inherent in my musical expression that was related to my language background. As I began to review the literature on the relationship between language and music, I discovered that the proposition has been mooted by thinkers throughout history, spanning a variety of disciplines. There is frequent mention of Plato and his belief in the resemblance between words and musical modes; that the modes could elicit responses of the spirit as strongly as words (trans. Griffith, 2000). From there, interest in music-language relations have appeared in various areas of knowledge, from differing perspectives. Darwin (1871) thought that the origin of the communicative abilities of our species may have been an intermediate form of modern language and music. Other philosophical theories connecting music with language include those of Girolamo Mei and

Vincenzo Galilei (see Gerbino and Fenlon, 2006), Jean-Jacques Rousseau (see Rousseau, 2009), and Ludwig Wittgenstein (see Guter, 2004).

More recent work concerning the relationship between language and music has been undertaken within and between the fields of philosophy, musicology, music psychology, linguistics, and cognitive science (Kania, 2013; Swain, 1997; Sloboda, 1986, 2004; Jentschke, 2016; see also the proceedings of the 9th Speech Prosody Conference, 2018). Within these fields, there are those who make analogies between the structural systems within language and music (Lerdahl and Jackendoff, 1983), those who theorise on the similarities between aural impressions of spoken languages and music (Abraham, 1974), those who believe that music *is* a language (Cooke, 1959), and those who seek, through scientific methods, cognitive parallels within the domains of music and language (Patel, 2008). I became particularly interested in studies that focus on speech production in relation to music (Palmer and Hutchins, 2006; Patel, 2008). One notable path of research involves the comparison of speech rhythm to the rhythm of instrumental melodies from composed music (see Patel and Daniele, 2003, 2006; Huron and Ollen, 2003). Researchers investigated the musical themes of composers from different cultural (and therefore language) backgrounds, discovering a plausible link between the speech rhythm of a culture's language and the rhythm of its music. However, these studies focus on observations of score-based, rather than live, performed music. On a similar path, although with a more generalised approach to speech, the work of Palmer and Hutchins (2006) involves relating the rhythmic and melodic properties of speech—termed 'prosody' in the discipline of linguistics—with musical rhythm and melody. They propose the concept of 'musical prosody', which views musically expressive elements of phrasing, *rubato*, and dynamic and articulatory variation in Western classical music performance as analogous to

the acoustical changes of frequency, duration, and intensity that form properties of prosody in speech. Unfortunately, this study makes no distinctions between different languages, never mind the native languages of performers.

While a significant portion of my own research was initially based upon (and continued with reference to) the above-mentioned studies, I sought to combine certain concepts and methods from different disciplines in order to answer my particular research questions. I was intrigued by the idea of undertaking an empirical study utilising methodology from linguistics, similar to that of Patel and Daniele (2003). From Palmer and Hutchins, I applied the concept of performed music as having analogous acoustical properties to those of speech; this provided a common framework within which to conduct my comparative, empirical investigation. Finally, I also made use of analytical techniques from current musicological research that specialises in the study of live recordings and live performance (see Cook, 2009 and Leech-Wilkinson, 2009; more detail on this in Chapter Three).⁵

1.4 Research Process and Questions

As outlined above, my review of the literature on the relationship between speech and music involved delving into a wide range of disciplines, including linguistics, cognitive science, musicology, and philosophy; also sub-disciplines, such as perception and production in speech and music performance studies. It had become increasingly evident as I discovered the scope of this literature that I needed to employ an interdisciplinary method

⁵ For analytical methods, reference was made to the Arts & Humanities Research Council Research Centre for the History and Analysis of Recorded Music.

of investigating my specific concerns. Of course, there was the danger—as there arguably is with all interdisciplinary work—of jeopardising a thorough and comprehensive grasp of one field in order to review several others. However, the fields of linguistics and music have a long-established shared interest, as well as a growing literature of correlative research; my task, therefore, had been to familiarise myself with the areas of overlap that have relevance for my research questions. My scope, at the beginning, was rather wide, but as I gained more perspective I was able to establish more specific areas of focus.

What's Missing in Current Research?

Starting from a broad spectrum and narrowing down to specialised literature provided me with the foundations for thinking critically about current research. I discovered that the wealth of literature on the speech-and-music connection is either theoretical (and, at times, highly impressionistic) or perceptual, from a linguistic or cognitive point of view. In both cases, it lacks any live musical contribution. Additionally, those very few studies that have privileged the musical perspective have been conducted by music psychologists. In these cases, performing musicians, if they are included at all, hold the status of necessary participant, with little contribution to the research process. Furthermore, while excerpts of scores or musical performances are used for comparative analysis, speech data are generally non-existent or simply referenced from other sources. In some studies, language is located with the umbrella concept of culture, with no reference to speech at all.

In general, few studies on production with regard to the speech and music relationship exist; in particular, there is a lack of comparative studies on speech production and musical performance. Most importantly, none of the existing studies grants agency to the performer as a researcher and a valuable contributor to the current scholarship. For

Doğantan-Dack (2008: 302), the lack of a 'performer's discourse', even within the musicological discipline of performance studies, is concerning. In Doğantan-Dack's (2008: 299) view, the 'embodied expert knowledge' of a performing musician is particularly valuable to current musicology as the ontological status of music shifts from a score-based conception to one that focuses on performance (see Cook, 2001, 2003, 2007; Clarke and Cook, 2004). Yet the experiences of performing musicians are rarely seen as having the same epistemological status and credibility as the theoretical and analytical literature of musicologists. This makes it all the more imperative for performing musicians to 'set aside their notorious image as inarticulate musicians, and fill in the epistemological gap' (Doğantan-Dack, 2008: 302) by articulating the perspective of performers in scholarly discourse. This became one of the chief aims of this present project: to conduct research from a performance perspective, so as to elevate the status of the thinking performer. There is a sense in which this, too, is a matter of locating a voice and allowing it to be heard. Ultimately, I recognised that the process of finding my own voice would involve both contribution towards further knowledge through research, as well as development as a musical performer. Thus, I decided to accompany this thesis with a concert performance, so that my voice can be heard through both mediums of language and music.

As my literature review progressed, I realised that certain questions, formed by my experiences as a performer, could not be answered by current research. A significant question is: 'Does a pianist's native language influence the musical prosody of her or his performance?' Since, thus far, there have been no studies focused on both the speech and musical performance of classical pianists from differing language backgrounds, I formulated an empirical experiment that combines methodologies from work on speech rhythm, rhythm in musical performance, cross-linguistic studies, and cross-cultural studies in music

performance. My work has also led to the confrontation of a wider, more personal question that has influenced and affected me throughout my musical life: 'Do I have *authority* as an East Asian musician to express Western classical music in the way that I can and do?' Of course, one could answer: 'Yes, because who is in a position to deny you that freedom?' Yet this answer poses another crucial question: who, or what, has denied me my sense of freedom, autonomy, and ultimately musical identity, to the extent that I felt compelled to devote several years of a research degree to searching for it? This is a central question that I sought to answer through the second stage of my journey. Along the way, I engaged anew with the controversial concept of authenticity, re-evaluated my conceptions about musical expression and the discourses that influenced their development, and made a self-examination of my own musical practice. I also came in contact with ethnographic research on the experiences of other East Asians within Western classical music. Reading about others' experiences and listening to those of my colleagues and friends, I realised that I was not alone in my doubts. Due to the deeply individual and charged nature of this investigation, I determined that an autoethnographic narrative of my own experiences would be most practical and illuminating.

Methodologies

To explore the question, 'Does a pianist's native language influence the musical prosody of her or his performance?', I designed an empirical experiment that studies both the performance and speech of classically-trained pianists. To interpret the results of the experimental study, to situate the results within the context of Western classical music performance, and to gain an understanding of how I was affected and transformed by my research, I utilised the method of autoethnographic narration. This helped connect my

personal and musical experiences during my research period with wider social and cultural issues. By making use of 'emotional recall' (Ellis and Bochner, 2000), I relived certain experiences and created stories from the thoughts and emotions that emerged after retrospection. This process revealed a side of myself that was both intense and vulnerable, one that was (to put it starkly) no longer content to be without a culture, an identity, and a 'voice'. Through this self, I confronted my own music-making by reflecting upon and describing my practice and musical experiences. I was also able, finally, to reconcile myself to addressing certain deeply personal and sensitive aspects of my life, revealing their profound effects upon the development of my self-identity.

As with all autoethnographic narration, this part of the present submission is not meant to be read as factual description of what had occurred during the last several years; rather, it was more important for me to 'convey the meanings that [I] attached to experience[s]' (Ellis, 1999: 674). In this sense, the stories that I created seek to situate readers within my own 'inner world'; I focused on 'evoking' emotional responses and allowing for dialogues and further discussions to be formed, rather than forcing information upon 'passive receivers' (Ellis and Bocher, 2000: 744). In the manner of philosophical dialogue, some of my reflections are voiced by two characters, *One* and *Two*. Both voices are my own, and are an extension of my inner dialogue as I reflected upon my own research. *One* is characterised as sceptical and concerned, while *Two* seeks to explain and clarify the issues raised by *One*. Rather than arrive at any kind of objective truth, these dialogues seek to show both my introspective process and trace its outcomes within the text.

Finally, the culmination of my research journey is a recital in which I revisit the Berg Piano Sonata and include several works of chamber music, namely the Piano Trio in C minor

by Felix Mendelssohn, the Violin and Piano Sonata in A major by Gabriel Fauré, and the Piano Quartet by Robert Schumann. This thesis does not include detailed descriptions of my practice of these pieces in the traditionally musicological sense, nor can I guarantee that all my expressive decisions derive directly from results of my research: live musical performances are malleable and never strict reproductions of what occurs in the practice room (Doğantan-Dack, 2012a). However, I believe in the importance of having my music serve as the final outcome of my research process; I began my journey as a performer, was motivated by my own practice, and through my work, was able to further understand the position of my music-making, both within a broader social and cultural context and within my own internal sense of self. Although the textual element of this submission seeks to communicate that process of understanding through words, I want my performance to be the final word: I want to convey that my voice is musical.

Experimental Research Focus

My experimental research investigates speech and music as they are produced and performed. As a professional musician, I am highly experienced in listening to and analysing music as it is performed by myself and others. Thus, I am interested in data that can be extracted and analysed from sound productions, and do not aim to go beyond that in this thesis. Inevitably, my work refers to perceptual research in the fields of both speech and music; I have also been influenced by relevant literature in philosophy, musicology, and music performance studies, and likewise in cultural studies. Most importantly, this thesis does not aim to form conclusions about specific cultures and their abilities in the performance of Western classical music: my interpretation of my data is certainly not meant to re-inscribe larger racial or cultural biases in Western music! Rather, it argues that

expression, whether through speech or music performance, is a highly individual and complex process. It is a process that has pretensions only to beget further exploration.

1.5 Outline of Thesis

In Chapter Two, I situate my research among relevant literature; this involves an overview of the language-and-music/language-versus-music debate, with a focus on the parallels between the two domains. Also in Chapter Two, I describe the comparable acoustic variables involved in both speech and music, and delve selectively into cross-domain and cross-cultural empirical research. Chapter Three presents my experimental design and elucidates the interdisciplinary methodology that I utilised for my quantitative study, while Chapter Four involves autoethnographic reflections upon my own practice. This is followed by Chapter Five, which consists of an overview of the cultural and socio-political background of East Asian classical musicians and their families. This chapter also includes snippets of collected interviews and statements from ethnological literature, as well as snapshots of my own background and experiences as a child of Chinese immigrants in the United States. Finally, Chapter Six reads as a personal reflection on my entire research journey; I address my current thinking about my own sense of identity, as well as on the transformation of my relationship with authority and expression. These reflections take place within recollections of my practice of the pieces featured in my final performance.

Stage One

Chapter Two: Literature Review

2.1 Introduction

This chapter situates my research within relevant literature from a variety of perspectives (musical, linguistic, and cognitive) and fields (musicology, music psychology, music philosophy, linguistics, cognitive science, and neuroscience). I begin with an anthropological survey of scholarship within musicology and music psychology on the relationship between language and music prior to the 21st century (Feld and Fox, 1994), followed by a review of current developments in language-and-music research in the fields of cognitive science and neuroscience. This necessitates a discussion of the debate on the plausibility of analogy between language and music, followed by examination of the commonalities between the two. At this point, I provide an overview of the concept of ‘musical prosody’, introduced by Palmer and Hutchins (2006) and developed from the parallels between language (speech prosody) and music (expressive musical performance). This determines the perspective of my own research: rather than taking a musicological view on the connection between language and textual music (i.e. the score), I focus on analysing the properties involved in the actual performance of music and production of speech. From here, I examine the concept of musical expression within scholarship of the Western classical tradition, highlighting the complexity of its definition. My decision to maintain a performance perspective requires a shift away from existing musicological and philosophical considerations on expression in music, towards empirical work within phonetic science on the specific properties involved in speech production that parallel those in expressive musical performance. This leads to the establishment of rhythm, in both speech and music,

as the central focus of my empirical research. In the subsequent section, I discuss the comparable acoustic variables involved in speech rhythm production and expressive musical performance, and select duration as the rhythmic parameter of focus for my experimental study in Chapter Three. The final section reviews and critiques a selection of current cross-domain and cross-cultural studies, highlighting a need for further ecologically-valid, performance-based comparative research on speech and music.

The vastness of my topic, along with the practical limitations of this thesis, require that I stay true to my own perspective, that of a performing musician. Therefore, the following overview of the literature is not exhaustive, but rather focuses on work that informs and challenges my particular views.

2.2 Literature on the Language and Music Relationship

The relationship between language and music has been a topic of interest throughout the ages and among various disciplines; scholars within musicology, linguistics, literary studies, philosophy, psychology, and more, have sought to address this topic. In recent years, cognitive science, neuroscience and evolutionary biology have contributed to the study of music and language as cognitive systems, spurred by the belief that comparisons of the two domains can lead to further understanding of the structural and functional properties of each, and ultimately reveal more about how the brain works (Rebuschat, et al., 2011).

Scholarship Prior to the 21st Century

An anthropological survey of scholarship on the language and music relationship prior to the 21st century reveals four main perspectives: ‘music as language’, ‘music in language’,

‘language in music’, and ‘language about music’ (Feld and Fox, 1994). The first of the perspectives, ‘music as language’, centres around the idea that music is analogous to language, and therefore investigation of it benefits from the analytical techniques and models of linguistic approaches (e.g. to syntax, morphology, and phonology). Work in this category can be further divided into two traditions, the first of which is founded on analogies between the distributional organisation of pitches in music (for studies on listeners’ sensitivities to distributional information in music, see Krumhansl, 1990; Laden, 1994; Oram and Cuddy, 1995) and phonetic organisation in language (Feld and Fox, 1994: 29; see Springer, 1956; Nettl, 1958; Nattiez, 1977; Jakobson, 1987; Houghton, 1984). The second tradition, firmly grounded in Western music theory, is founded in parallels between the harmonic, metrical, or motivic organisation of musical works and the syntactic organisation of language. Work of this tradition has produced musical grammars and formal descriptions of phrasal, harmonic, and metric syntax of musical pieces and styles (Feld and Fox, 1994: 29; Lerdahl and Jackendoff, 1983; Sundberg and Lindblom, 1976; Raffman, 1993; Ruwet, 1967). Critiques of the ‘music as language’ perspective have focused on its bias toward metrically-regular, hierarchical, and tonal samples of Western art music, as well as its emphasis on ‘macro-syntactic’ (melody, rhythm, tonality, mode) rather than the ‘nuanced, micro-syntactic’ dimensions of music (pitch, texture, tempo, dynamics; Feld and Fox, 30). Furthermore, scholarship within this category is largely ‘score-centric’, focusing on the score as musical product, rather than elements involved in the process of music-making (i.e. performance).

The second perspective, ‘music in language’, involves work on discourse prosodics with a focus on prosody and paralanguage (e.g. tempo, dynamics, and voice quality) in speech and verbal art. Within this category, linguists and literary theorists have targeted the

musical dimensions of spoken discourse, i.e. the suprasegmental elements of stress, rhythm, meter, pausing, voice quality, and timbre (Evans and Clynes, 1986; Jackendoff, 1989; Kiparsky and Youmans, 1989; and Woodbury, 1992 for speech; see Hinton, 1990; Sherzer and Woodbury, 1987; Tedlock, 1983; and Woodbury, 1987 for verbal art). This perspective has contributed a variety of empirical work on discourse prosodics and is significant to the development of further-integrated research on language and music within linguistics.

‘Language in music’ involves the study of texted vocal music (Lomax, 1968; Boswell, 1977; and Yung, 1983), speech surrogates (i.e. drum and whistle languages; see Stern, 1957 and Umiker, 1974), and performative elements such as chant, recitative, *sprechgesang* and *Sprechstimme* (see Feld, 1989; Herndon, 1974; and List, 1963). The last perspective, ‘language about music’, centres on the verbal discourse of musical experiences (see Kingsbury, 1988 for discourse within the American music conservatory; see Crafts et al., 1993 for interviews conducted on American musical experiences).

Musicology and Music Psychology

Clarke’s (1989) overview and critique of literature within musicology and music psychology describe several contributions toward the use of linguistic models to describe music. Within musicology, Cooper and Meyer’s (1960) work on rhythmic structure in music employed a system of poetic feet to describe five general rhythmic types within the Western classical canon; Cooke (1959) developed a lexicon of musical expression based on the concept of language as a system that communicates emotional states between the composer and the listeners; lastly, Lerdahl and Jackendoff (1983) developed links between musical properties and prosodic features of language by basing their Generative Theory of Tonal Music

(hereafter GTTM) on Noam Chomsky's generative transformational linguistics (Chomsky, 2002). Clarke finds fundamental issues within these studies; in the first, the concern is related to the complexity of rhythm in speech and music and what he considers to be incompatibilities between the temporal properties of each (Clarke, 1989: 12; a detailed discussion of rhythm will be provided in a later section). The work of Cooke (1959) is deemed problematic due to its 'naïve' approach to textual music, the restricted repertoire on which the lexicon of musical expression is based, its lesser view of instrumental music compared with vocal music, and its neglect of the performer within the communicative model of composer as 'sender', the piece as the 'channel', and the listener as the 'receiver' (Clarke, 1989: 13). Clarke's difficulties with the work of Lerdahl and Jackendoff relate to their insistence on music as 'pure structure... not tied down to specific meanings and functions' (Lerdahl and Jackendoff, 1983: 9), which avoids engagement in context, signs, and styles that are inevitably present in the practice of musical analysis (not to mention performance; Clarke, 1989: 14). Also briefly mentioned is Bernstein's (1976) attempt at establishing direct links between music and Chomskyan generative theory in various ways (e.g. 'musical phonology', 'musical syntax', and 'musical semantics'); these links have been criticised as intuitive and lacking in empirical methodology, as well as a firm grasp of linguistic frameworks (Jackendoff, 1977). However, Bernstein's theories seem to have had lasting effects upon music scholarship, particularly as a predecessor for Lerdahl and Jackendoff's GTTM (Akbar, 2012: 123).

The work of Joseph Swain (1997), continuing that of Cooke, Bernstein, and Lerdahl and Jackendoff, contributes toward the concept of 'musical language' as a parallel model to spoken language by emphasising *context* and its significance within the perception of the

acoustic signal in both music and speech. Swain's theory addresses the concepts of 'phonology', 'syntax', and 'semantics' in music with regards to context, explaining that

sounds, whether phonemic or tonal, are never understood as isolated events that are strung together to make speech and music. Rather, every sound is heard in a contextual background, so that the listener takes an active and essential role not only in figuring out the grammatical function of the sound but in the very identification of the sound (Swain, 1997: 15).

Working towards a 'musical syntax', Swain states that 'syntax in music is more than a system that controls information by organising pitches and durations. Just as it mediates expressed relationships in natural languages, syntax mediates the relation of tension and resolution in musical languages' (Swain, 1997: 28). From this, he explains that the difference between the 'musical languages' of, say, Josquin des Prez and Johann Sebastian Bach may not only have to do with the respective means of organisation and hierarchical structuring, but also in the tensions and resolutions that they create (Swain, 1997: 34). As for 'musical semantics', Swain deals with the polarising concepts of music as having either no meaning and referring only to itself (Stravinsky, 1998) or as expressive of different emotions (Kivy, 1980 and Langer, 1953) by focusing on 'semantic range' in both spoken language and music. According to Swain, context allows for a broadening out or narrowing in of the many possible meanings of a certain word or a passage of music. Swain's work contributes to the on-going debate over the analogy of music and language, a topic with competing theories which will be discussed in detail following a brief overview of research within music psychology.

Scholars prior to the 21st century have conducted empirical studies on perception, investigating the existence of rule-based perceptual principles in Western classical music, as well as studying emotion and meaning in music. Sloboda (1976) studied pianists' sight-reading of conventional but unknown pieces in which notational errors were deliberately introduced; results demonstrate that the pianists both failed to notice the errors and unconsciously corrected them in their playing, suggesting the existence of tacit knowledge of certain stylistic norms or principles.⁶ The work of Diana Deutsch (1999) suggests that memory for pitch and low-level pitch relationships (e.g. intervals and chords) is based on a number of highly specialised systems, while higher level pitch information (e.g. contour) is retained based on hierarchies. One interesting study on the perception of two-tone pairs that are related by a half-octave (i.e. tritone) demonstrates that listeners with different language backgrounds (from California and Southern England) disagreed as to whether a given pair formed an ascending or descending pattern; when the Californian group tended to hear the pattern as ascending, the Southern English group tended to hear it as descending (Deutsch, 1991). This, along with an earlier study (Deutsch, et al., 1990) which shows a correlation between listeners' perception of this 'tritone paradox' and the pitch-range of their spontaneous speaking voice, suggests an interacting relationship between speech production and perception of this musical pattern. Empirical work in music psychology on emotion and meaning in music includes contributions by Gabrielsson (1973), Imberty (1975), Gabriel (1978) and Clynes (1982). Dowling and Harwood (1986) took a

⁶ This study also included reading of verbal texts which included spelling errors; results showed similarities between both music and language reading regarding error detection.

semiotic approach to explaining musical meaning as consisting of a mixture of intentional, extensional, rhetorical, and pragmatic components (see Clarke, 1989: 20).

Scholarship of the 21st Century: A Debate

Turning towards more current literature involves discussion of the debate over language and music's comparability. This subject is divided into two views, one emphasising the differences between the two, and the other convinced of the similarities. Beginning with the differences, one argument is that language involves the forming of sound patterns into words, which are then used in grammatical categories (nouns, verbs, and adjectives) and functions (subject, direct object, indirect object), while such categories and functions are lacking in music. Language and music are seen to have different ecological functions in human life; words in language convey 'propositional' and 'conceptual' thought (Jackendoff, 2009) and music conveys affect. While language is also able to evoke emotion, and can be used affectively, music is limited in its propositional function. Furthermore, a central characteristic of all known languages is 'meaning specificity', which allows for reference to the 'extra-individual world' (objects which can be perceived and thus verified, falsified, or proven to exist by different individuals; Koelsch, 2013: 171). Music seems to refer more to sensations felt within an individual and is also 'far less specific' than language in terms of meaning (Slevc and Patel, 2011: 110). 'Translatability' is another characteristic associated much more with language than with music (Patel, 2008: 301); a language can be translated more or less directly into another. An example offered by Jackendoff (2009) is that Japanese can be translated into Swahili or Quechua, but Japanese cannot be translated into raga, rock and roll, or Japanese *gagaku*.

These differences have been challenged, or rather approached, from various sources. One such source is Koelsch (2013: 169), who offers an alternative view of the language-and-music relationship: he proposes a ‘music-language’ continuum, which acknowledges that while language and music have differing characteristics, they do not exist in a ‘clear-cut dichotomy’. Instead, concepts such as ‘propositional’ and ‘non-propositional’ semantics, ‘meaning specificity’, and even ‘translatability’ exist in a continuous system with a language pole on one end and a music pole on the other. This accounts for the instances of imprecision of quantifiers, modals, and connectives in language and the ability of Western tonal music to convey quantifiers (‘some’ and ‘all’ shown by variation of musical texture), modals (‘must’ can be shown with music of strong intention), and connectives (structural relationships between musical elements within a composition). A continuum of ‘meaning specificity’ allows for ambiguity of linguistic meaning as well as the ability of affective prosody to alter the meaning of uttered words. In Western classical music, the *leitmotif* (a short, recurring musical phrase that is associated with a particular place, person, or idea, above all in opera) functions similarly to specified meaning in language. As for ‘translatability’, there are issues related to the translation of text from one language to another, while recent work has approached the ‘translation’ of music across instruments, genres, and styles through machine-learning methods (see the Universal Music Translation Network developed by Mor et al., 2018).

A significant amount of research in cognitive science and neuroscience has also contributed to the language and music debate. Music and language have traditionally been treated as two separate psychological faculties due to the theory of laterisation of music and speech functions within the brain: speech functions were thought to be localised in the left hemisphere and musical functions in the right hemisphere (Bever and Chiarello, 1974).

This view has been challenged in recent years due to the advancement of modern brain imaging techniques and the improvement of neurophysiological measures. Furthermore, studies on the relationship between tonal language expertise and musical pitch perception support a cross-domain influence of language experience upon pitch perception, e.g. Mandarin speakers demonstrated higher sensitivity to subtle pitch changes and interval distances compared to a control group (Guiliano et al., 2011). Congenital amusics (as the terminology has it) also demonstrated impairments of pitch-processing in speech, specifically in tonal languages (Tillmann et al., 2011).

Research also suggests that musical expertise interacts with language functions (Ott et al., 2011; Gordon et al., 2011; Strait and Kraus, 2011) and that it may be beneficial for phonetic perception (Patel, 2011). In an investigation on the effect of a musical chord's tonal function on syntactic and semantic processing, Hoch et al. (2011) revealed interactive effects between music-syntactic and linguistic-syntactic processing, as well as music-syntactic and linguistic-semantic processing. Patel (2012) compared the event-related brain potentials (ERPs) of 15 musically educated adults, elicited by syntactic incongruities in language and music. Results revealed that both linguistic and musical incongruities elicited positivities that were statistically indistinguishable. These studies suggest that there may be shared neural resources for structural integration processes and sequencing in the domains of music and language. These studies are not uncontroversial: the theory of modularity or domain-specificity (Peretz and Coltheart, 2003; Peretz, 2009) argues that language (speech) and music have distinct, functional specialisations within the brain. In response to Patel's (2008) theory of shared neural resources, Peretz (2008) speculates on the possibility of domain-specificity *emerging from* the shared resources; in other words, music and language

may have a number of components that relate to a general-purpose mechanism, but become highly specialised during the activity of processing.

Evolution

On the topic of debate, research on the evolution of music and language is likewise divided into two views: one, that music and language come from a single, primitive form, hypothesised as either a ‘musilanguage’ (Brown, 2000) or a ‘proto-language’ (Wray, 1998, 2000; Arbib, 2002, 2005), or two, that music and language evolved parallel to each other as ‘independent evolutionary specialisations of primate communication’ (Jackendoff, 2009: 198). Mithen (2005) believes that, on the basis of what they share, it is unlikely that language and music have independent evolutionary histories: both are universal features of human society; have three modes of expression (vocal, gestural, and written); are hierarchical systems; and can be used to express emotion. However, Jackendoff (2009: 202) argues that certain shared features may have evolved from ‘older functions’ that could have been appropriated by both language and music independently. Such features include the conveyance of affect, as well as ‘downward’ endings of both spoken and musical phrase contours; both could have been appropriated from mammalian call systems. A further argument involves experimental evidence suggesting that the hierarchical structures of both music and language are integrated by the Broca’s area in the brain (Patel, 2003, 2008). Jackendoff (2009) speculates that this area also integrates complex action structures, suggesting that language and music may have evolved, independently, from a function that governs the integration and execution of complex action.

This review represents a far-from-comprehensive consideration of the diverse literature surrounding the relationship between language and music. As it shows, there

continues to be much debate as to the relationship between language and music, both conceptually and as cognitive systems. For the purpose of this thesis, however, I will focus on surveying work that emphasises the commonalities between language and music, as my research questions locate themselves within the connections that exist between the two domains.

2.3 The Connection between Speech and Musical Expression in Performance

Regardless of the differences in ecological function and formal structure, and of the theory of modularity, it can surely be agreed that both language and music involve the production of sound sequences that vary in timing, intensity, and frequency, as well as particular types of organisation, coordination, and grouping. Cognitively, both domains require substantial memory capacity to store representations (words in language; pitches, melodies and harmonies in music) as well as the ability to utilise stored representations within working memory via rules and ‘structural schemata’ to execute complex actions such as verbal communication and musical performance (Jackendoff, 2009). Additionally, both domains support perceptual expectation (the ability to predict or expect what is to come; Patel, 2012; Patel and Morgan, 2016).

There is currently a growing body of research dedicated to the commonalities between language and music of the Western classical tradition: Patel’s (2008) work not only includes cognitive investigations of the relationship between the two domains, but also contributes to interdisciplinary, comparative research on speech and instrumental classical music. A particularly relevant group of studies involves the investigation of musical themes of composers from different cultural (and therefore language) backgrounds (Patel and Danielle, 2003; Patel et al., 2006; see also Huron and Ollen, 2003). Results show a plausible

link between the speech rhythm of a culture's language and the rhythm of its music. While these studies focus on comparing speech with textual data, i.e. the notated durations prescribed by the composer, they have been influential also with respect to further cross-domain research on data extracted from live musical performances (e.g. Carpenter and Levitt, 2016; McGowan and Levitt, 2011; others to be discussed below). Another interesting area of work involves the relation between the rhythmic and melodic properties of speech known as 'prosody' with rhythm and melody in musical performance. Palmer and Hutchins (2006) propose the idea of 'musical prosody', which views musically expressive elements of phrasing, *rubato*, and dynamic and articulatory variation in Western classical music performance as analogous to the acoustical changes of frequency, duration, and intensity that form properties of prosody in speech.

Musical Prosody

Much of relevant research has focused on comprehension and perception within the language and music domains. Combinative studies on speech production and musical performance are rare, but becoming less so; in recent years, scholars in music psychology have theorised and collected experimental data on the connection between rhythm in speech and rhythm in expressive musical performance. Caroline Palmer (1989) and Palmer and Hutchins (2006) have investigated performers' expressive nuances, relating such individualistic phenomenon to speech prosody. What in linguistics is known as prosody is systematic within each language and distinct from other structural levels of linguistic analysis (Beckman, 1996; Pierrehumbert, 1999). Performers manipulate music in a similar fashion: while still aware of the categorical pitch and duration values determined by the composer, they subtly alter the frequency (pitch), duration (time), amplitude (dynamics),

and timbre (harmonic spectrum) to produce musical expression. Palmer and Hutchins (2006) introduce the term 'musical prosody' to address these expressive nuances, which also appear to be systematic. For example, studies have shown that musicians utilise 'prosodic' features of slowing down at phrase endings during expressive performances, even where the composer did not indicate gradual lengthening or extension of final note values (Palmer, 1989). Also, musicians' prosodic choices of timing and intensity change systematically depending on the musical context (Palmer and Hutchins, 2006); it was observed that performances of melodic excerpts showed greater phrase-final lengthening when isolated and not embedded in a longer sequence. Also, performances of the same melodic excerpt embedded in different metrical contexts (sequences in triple and quadruple time, for instance) exhibit different prosodic features (Palmer et al., 2001). It is likewise possible for listeners to recognise prosodic cues in musical performance; a study using experienced and inexperienced listeners revealed that both groups could distinguish between musical phrases that differed only in expressive nuances. Furthermore, listeners could recognise when expressive nuances did not match the rhythmic context of the excerpts onto which they were superimposed (Palmer et al., 1999). These findings suggest that prosodic cues in music performance are part of listeners' memory for melodies and that such cues can enable recognition of auditory events in music just as in speech.

At the centre of this work is the interesting question of whether 'musical prosody' derives entirely from the individualistic interpretations of performers, or if the expressive features are somehow 'mapped according to cognitive principles shared across performers and listeners' (Palmer and Hutchins, 2006: 8). According to Palmer and Hutchins, both are involved; performers make individualistic decisions concerning segmentation (e.g. marking phrases or highlighting individual parts within simultaneous music, known as 'voice leading')

and prominence (e.g. marking elements within a phrase that are metrically or structurally important), yet certain 'prosodic cues' (e.g. further shortening of lengths of short note durations, such as a quaver; lengthening of long note durations, such as a minim; phrase-final lengthening; and deceleration of performance tempo at the end of a musical performance) are general enough to be expressed by rule-based models (Sundberg, et al., 1983; Sundberg, 1989; Todd, 1985, 1995; Sundberg and Verillo, 1980). These rules have been based on a variety of sources, including the musical intuitions of a trained performer (Sundberg et al., 1983, 1989; Sundberg and Frydèn, 1985) or on Lerdahl and Jackendoff's GTTM (see Sundberg and Verillo, 1980; Todd, 1985). Such models can generate performances that include subtle manipulations of pitch, duration, and intensity based on the structural information of the notated score. These rule-based performances, when tested for preference against computer-generated performances that did not apply the rules, were favoured by both musicians and non-musicians (Sundberg, et al., 1991). This implies that there are certain general prosodic features in performance, accepted by listeners, that allow for the perception of musical structure by way of segmentation and prominence. However, the fact remains that different performers of the same composition, motivated by differing understanding of the musical structure and hence differing interpretative intentions, apply different prosodic cues (Kendall and Carterette, 1990; Palmer, 1989). Palmer's (1989) experiment, using Mozart's piano sonata in A major K. 331 performed by six different pianists, demonstrates that the use of chord asynchronies, *rubato*, and articulation (staccato and legato patterns) varied across performers and corresponded to their individually-written interpretations (labels marked upon the score) of structural boundaries and voice leading.

Before any further attempts to address the coexistence of both individualistic performer-based interpretations and systematic expressive features based on ‘cognitive principles’ can be made, it is necessary to acknowledge the existing scholarship on the topic of musical expression in Western classical music.

Musical Expression and Interpretation

Theorising the involvement of both individualistic and systematic expressive cues necessitates delving into literature on expressivity in Western classical music: this concept is complex and confounding in its definition due to the range of historical and philosophical perspectives from which it has been viewed. Prior to 1800, music’s expressiveness was thought to inhere in, variously, its ability to imitate both animate and inanimate nature, its ability to imitate the passions, and its direct arousal of the passions within the listener. After 1800, music was regarded as the self-realisation of the spirit and all the emotions surrounding the Soul (Hegel, 1975) or of the Will (Schopenhauer, 1958). By the 20th century, following the apogee of Expressionism, music could be heard as non-expressive (Stravinsky, 1947). A brief overview of such perspectives prior to the 21st century reveals a significant focus upon the composer, the work, and the listener, and a general lack of interest upon the performer and/or the performance (see Baker et al., 2001; as well as Goehr et al., 2001); the only mention of expression in relation to the specifics of actual performance can be found as a hypothetical situation in which a piano teacher may ask a pupil to “‘put in expression”, i.e. to play a piece with a certain articulation, tempo, and phrasing’ (Baker et al., 2001: 1). Philosophical considerations of ‘performance’ by the end of the 20th century have typically focused the concept of ‘authenticity’, which places importance upon the composer and the work, rather than on the performer or the performance itself (with exceptions in jazz and

electronic music; Davies, 2001). As for the topics of ‘expressiveness’ and ‘understanding’, the areas of investigation have typically been emotion (see Raffman, 1993; Langer, 1942; Goodman, 1968; and Kivy, 1989) and arousal (see Ridley, 1995; Matravers, 1998), as well as of the listener’s musical understanding (see Meyer, 1956; Kivy, 1990; Davies, 1994; and Levinson, 1997). A similar survey of work within musicology reveals little mention of performance and the role of the performer except for a brief comment on growing postmodernist interest in seeking to understand musical expression ‘independent from the structure’ (Duckles and Pasler, 2001: 7). Fortunately, this ‘growing interest’ has recognised the predominance of literature on ‘music as writing’, as opposed to ‘music as performance’ (see Cook, 2013; for a philosophical perspective, see Godlovitch, 1998 and Kivy, 1995). The scholarship collected in Rink (1996, 2002), combined with the developing field of artistic practice as research (see Doğantan-Dack, 2015) demonstrate the shifting focus onto performance and practice in the study of musical expression (see Fabian, et al, 2014).⁷

A term that may often be synonymous with the musical expression of the performer is ‘interpretation’. However, Davies and Sadie’s discussion of this concept (2001) again places the composer at the centre, as the interpreter (i.e. performer) may convey her or his own idea of the music through performance, but only according to the will of the composer through the score. The leeway permitted to the performer to form her or his own view of the best way to convey the score is judged based on how closely the instructions and understandings of the composer are met. Also, mentioned several times is the notion that most musical works are capable of inspiring ‘multiple revealing, aesthetically rewarding interpretations’; however, this is again restricted by the sense that, although there is no

⁷ Further references are included in the section on expressive timing research below.

single interpretation that is deemed 'correct', there are those deemed 'incorrect' if they can be shown to run contrary to the aforementioned instructions and understandings of the composer (Davies and Sadie, 2001: 2). From the perspective of a performing musician, there seems to be a gap between this understanding of interpretation and the act of expressive performance. The missing link relates to further understanding of 'how' an interpretation is actively made during performance, and whether this process is as exclusively concerned with the will of the composer and the score as is assumed in the literature. Davies and Sadie (2001: 2) briefly mention the kinds of interpretative decisions involved in the 'full sonic detail of any actual performance' as involving both micro-level ('affecting subtleties of attack, intonation, phrasing, dynamics, note lengths, and the like') and macro-level ('concerning the overall articulation of the form, the expressive pattern etc'). However, precisely how these decisions are actualised by a performer is left undetailed. Interestingly, such emphasis on expression as related to the compositional, textual, and structural within musical academia prior to the 21st century mirrors that of linguistics, where the focus had been more on structural and grammatical prosody (e.g. tone, phonologically-determined duration, and intonational contours such as those indicating questions) than on 'performance' prosody individual to the speaker.

Universal and Biological Principles in Music and Speech: Emotion?

This section provides a brief discussion of literature treating certain 'cognitive principles' related to emotion that may be the foundation for systematic expressive cues in 'musical prosody'. Many of what may be considered the universal principles guiding expressive musical performance involve affect or emotion. Some view music (specifically vocal; instrumental by association) as physiologically related to emotional vocalisations (Juslin and

Laukka, 2003; for a study on intensity in affective speech and instrumental performance, see Liu and Xu, 2015). The voice, which becomes physiologically changed due to emotional arousal, produces certain acoustic patterns of respiration, phonation, and articulation that are recognisable to listeners. Music, it is believed, communicates emotions by using specific patterns of acoustic cues derived from vocal expressions of emotion (Juslin, 1998). Of course, this is not a complete explanation of how music elicits emotions within the listener, since elements such as harmony and tonality also contribute to emotional or affective responses. It is interesting, however, to note that parallel concepts exist in linguistics regarding affective and paralinguistic prosody. Gussenhoven (2001) discusses 'biological codes' of speech production and pitch, based on two biological conditions: one, production of speech involves energy and is tied to the breathing process, hence the creation of phrases; and two, humans' speech organs vary in size and weight, leading to different fundamental frequencies (see Ohala, 1983, 1984, 1994 for the 'Frequency code'). According to Gussenhoven, the 'Effort code', 'Production code', and 'Frequency code' explain the existence of certain universal prosodic cues. The 'Effort code' states that more energy is required for more precise articulation, as well as numerous and more varied pitch movements (this relates to affective interpretations such as 'surprise', 'authority', and 'pleasantness'). The 'Production code' states that the use of energy in conjunction with the exhalation phase of the breathing process results in breath groups (see Lieberman, 1967) and that utterances begin with high pitch and end with low pitch (this relates to informational interpretations such as questions or continuations and endings). The 'Frequency code' is often used for affective expression (high pitch for femininity and vulnerability, low pitch for masculinity and assertiveness). Affective prosody also involves

intensity (see Zimmermann et al, 2013 for an evolutionary theory of affective prosody and its relationship with music and non-verbal communication in non-human mammals).

The question of whether the systematic cues within ‘musical prosody’ can be explained by certain universal and biological principles that exist in speech, by way of emotional responses, is beyond the scope of this thesis. However, investigating the properties of expression involved in both speech and music may shed light on how it functions and how it may be learned and developed by an individual. Already there exists work that supports the concept of expression as trainable (see Clarke, 1993; Repp, 2000; Lisboa, 2005; and Timmers and Sadakata, 2014), but more research is necessary. There is also the possibility that expression is a skill or propensity that functions within both music and language domains with significant overlap. Thus, an individual may have a wide expressive range in both speech and musical performance, irrespective of language background. In this case, there is still a link between linguistic and musical expression, however linguistic ability would not be a determining factor in the use of expressive devices during musical performance. Palmer and Hutchins (2006: 40) pose some interesting questions: ‘Are musical genres, like languages, points on a prosodic continuum that is bounded by general perceptual principles of rhythm, grouping, and prominence?’ ‘Do individual differences in music or language abilities arise from differences in sensitivity to prosody?’ In the following section, I discuss and compare the principles of rhythm, grouping, and prominence within speech and music in detail. My own exploration of the second question, through empirical experimentation, will be found in Chapter Three.

2.4 Rhythm

The concept of rhythm exists in both domains, i.e. it is an integral part of speech prosody as well as musical expression. Although there are certain differences in the way rhythm functions in each, the parallels are significant enough for scholars to draw comparisons (see Hawkins et al., 2013; Hawkins, 2014). Rhythm, itself, is conceptually complicated and multi-contextual; there is no universally accepted definition, and it may be applied not only to auditory phenomena but also visual (e.g. in the visual arts and architecture) and within biological systems (e.g. circadian rhythms and brain oscillations). In speech and music, it can be understood as any systematic temporal, accentual (which involves pitch and intensity), and phrasal patterning of sound. Theories of speech rhythm have been varied and controversial due to the multiplicity of the elements involved, as well as the conflicting methodologies used (see Post and Payne, 2018 for an overview of speech rhythm theories). In music, researchers have shown a growing concern for ecological validity regarding psychological measurement and experimental design. Studies have shown, for example, that human performers never precisely produce ratios of 2:1, 3:1, and 3:2 as are indicated by Western rhythmic notation (see London, 2001). Also, listeners do not expect to hear such ratios in performance (Repp, 1995). ‘Real-time’ concert performance data rather than notation-based data reveals a more complex but richer picture of rhythm in perception and performance. Likewise in linguistics, earlier rather simplistic models of speech rhythm being centrally coordinated have in recent years given way to a more complex understanding of the multiple structures that may contribute to the ‘rhythm percept’ (see Nolan and Jeon, 2014; White, 2014), combining multiple structural factors of speech (such as lexical stress, phrase boundaries, and lexically-contrastive duration) and recognising multiple phonetic parameters through which these are conveyed (spectral information, duration, pitch, and

intensity), including the impact of performance factors (affect, speech rate, speaker idiosyncrasies). It is believed that an understanding of the multiplicity of systems and parameters involved can lead to a fuller understanding of the percept of rhythm in a speech utterance (Post and Payne, 2018).

As mentioned above, there are many commonalities as well as significant differences across the domains of speech and music. One difference, specifically regarding Western classical music and speech rhythm, lies in the relationship between design and acquisition. In Western classical music, rhythm (both large-scale, i.e. metre, and small-scale, i.e. rhythmic figures and patterns) is usually predetermined by the composer and notated in the score. Although the performer can manipulate timing in subtle ways, the resulting rhythm(s) should still be recognised as matching the indications of the composer, e.g. time signature and relative note durations. Rhythm, in this case, functions as both the temporal 'glue' of a piece of music, binding together melodic and harmonic events within a cohesive, temporal experience, as well as contrastive stimuli that are used to show musical character, imply physical movement, and/or engage the interest of the listener. The performer then has two responsibilities: to follow (to a certain extent) the rhythmic indications left by the composer, and deliberately but subtly to deviate from the regularity of the rhythmic notation for the sake of artistic interpretation. In most instances of natural speech, rhythm is not a conscious design of the speaker, but rather a perceptual outcome of the sound patterns in a language and performative shaping of an utterance within the limits of those patterns. Any given language seems to have certain rhythmic characteristics, contributed in part by the fundamental structural properties of the language that cannot be changed by the speaker. These structural properties are acquired in infancy, and their function has been theorised in several ways: Allen and Hawkins (1980) suggest that it is to do with the predictability of

shapes in utterances, which aids speech comprehension and child speech development; it is also thought to aid speech processing and talk-in-interaction (i.e. conversation; Hawkins, et al. 2013). Schreuder (2006) believes that rhythm is the means by which speech elements are carried and coherently packaged such as to transmit the linguistic message of an utterance.

Another difference between speech rhythm and musical rhythm involves the concept of isochrony. The view that speech rhythm cannot be compared with musical rhythm because of the lack of isochrony or periodicity in speech has both supporters and opponents. In much of Western classical music, rhythm can be understood as that ‘which allows for foot-tapping or hand clapping’, ‘with the organisation of notes into bars which are isochronous’ (Nolan and Jeon, 2014: 2). There is music, however, that purposefully rejects or deviates from this concept by use of compositional techniques such as syncopation, mixed metre, polytempi, or free time (a type of metre that is free from musical pulse and time signature). Examples include Charles Ives’ Piano Sonata No. 2, ‘Concord, Mass., 1840-1860’ (free time); Ives’ *The Unanswered Question* (polytempi); and *Old and Lost Rivers* by Tobias Picker (mixed metre). Furthermore, according to London (2001), the concept of metre or pulse regularity only came to existence in Western music around the turn of the 17th century alongside the development of ‘modern’ Western musical notation (which came from the need for a representation of the musical structure of a piece to be performed by different performers—even those without prior knowledge of the musical structure and manner of performance—at various occasions; London, 2001: 24). The general term of ‘rhythm’, understood by the musician or musicologist, ‘signifies a wide variety of possible patterns of musical duration, both regular and irregular’ (London, 2001: 2). The adjective ‘rhythmic’, used in music, commonly refers to a ‘metrically regular series of events’, however London (2001) notes that irregular rhythms can occur within the context of a

regular metre and that some metres may not require regular or even patterns of duration (e.g. Bartok's 'Bulgarian' rhythms). Thus, the characteristics of speech rhythm coined by Nolan and Jeon, such as 'antirhythmic' (relating to languages that do not have regular, isochronous lexical stress) and 'arhythmic' (those languages without clear lexical stress at all), may likewise pertain to music. Therefore, not all Western classical music is 'fundamentally rhythmic' (Nolan and Jeon, 2014: 8), at least not if we limit our definition of rhythm in terms of isochrony.

There are suggestive similarities between rhythm in both domains that have informed my research questions and experimental process. Fundamentally, the 'experience', or 'percept' of rhythm may be understood as the 'the temporal [or spatial] organisation of a sequence of similar events or objects, and of other parameters defining the relationship between these objects/events, i.e. contrastive properties' (Post and Payne, 2018: 7). In the case of both speech and music, this organising and contrasting of rhythmic events is signalled by sound, or rather more specifically by the acoustical parameters of duration, intensity, and fundamental frequency. Furthermore, if isochrony is viewed as only one of several possible types of grouping units in both speech rhythm and musical rhythm, then the possibility for comparing rhythm between the two domains (as experienced in speech and musical performance) is greater and potentially more fruitful. Another similarity between domains involves the alternation of elements. Several researchers have identified two distinct but relatable components of speech rhythm, termed 'periodic' (see Couper-Kuhlen, 1986: 51) or 'coordinative' rhythm, and 'contrastive' rhythm (White, et al., 2012: 1; White, 2014). Coordinative (or periodic) rhythm refers to a regular recurring pattern of events, i.e. isochrony, and contrastive rhythm is the alternation between strong and weak elements. White (2014) notes that while periodicity implies contrastive rhythm, the reverse

does not apply. Contrastive rhythm, essentially, is about sequentiality, i.e. elements that can be compared to those surrounding them. Contrastive rhythm dominates current theories of speech rhythm, while periodicity has more or less been refuted. Certain languages, such as English, use stress at both the lexical and phrasal level to denote contrast; however stress is not the only marker of alternation; Korean, for example, seems to use boundary cues (involving pitch and duration) to mark events that alternate (see Jeon, 2011; for perceptual studies, see Jeon and Nolan, 2013). Essentially, if one holds the broader view that rhythm in speech and music is an experience of contrastive events that are both sequential and alternating, then it is possible to conduct cross-domain studies on the elements involved in such events, studies that could potentially lead to further understanding of the relationship between the two domains.

Intensity

Within the multi-parametric model of speech rhythm, intensity is one of four parameters potentially involved in the rhythm percept, the others being duration, fundamental frequency, and spectral information (see Lee and Todd, 2004 for intensity variability; Arvaniti, 2009 for grouping and prominence patterns). Likewise, in music, intensity (or dynamics) is a rhythmic parameter just as duration and pitch are. It is necessary at this point to differentiate between the intensity variation involved in expressive musical performance and the dynamic variation involved in musical composition that functions interdependently with rhythm and pitch to create musical meaning and structure (Thiemel, 2001). In musical rhythm, intensity variation works together with pitch and durational variation to create patterns of prominence and grouping associated with the rhythm percept. Prominence is particularly important for rhythmic organisation (see Clarke, 1999 for a review of rhythm

studies). Intensity and duration are highly interconnected and studies have shown that listeners tend to hear intensity and/or pitch differences of evenly-spaced notes as durational differences. Also, durational differences may be heard as intensity differences, and the intensity difference of one note may affect the durational judgements of other notes (London, 2001; cf. Handel, 1989). Similarly, intensity, duration, and fundamental frequency combine in speech to mark auditory prominence and boundary cues, along with being involved in performance factors such as affect and speaker idiosyncrasies. Studies suggest that an increase in both duration and intensity are involved in the perception of prominence, while an increase in duration and decrease in intensity are perceived at pre-boundary positions (see Mo, 2008 for a study on American-English).

2.5 Comparable Acoustic Variables in Speech and Music

As mentioned above, both speech and music involve the production of sounds that vary in timing, intensity, and frequency (pitch). These variations result in structural features of prominence and boundaries, which allow listeners to perceive and process a continuous flow of sound as a series of meaningful events. Prominence involves the highlighting of elements of relative importance, however defined, and boundaries occur when continuous sound is parsed or segmented into smaller, more perceptible units. Speech prosody also involves other meaningful performance features such as speech rate and affective use of intensity, voice quality, and pitch. Likewise, musical expression involves tempo, timbre, and subtle pitch deviations (for non-keyboard instruments). For the sake of focus, my study examines only one of the acoustic parameters: duration, which is involved in the signalling of both structural features of prominence and boundaries in speech and their equivalence in music, as well as factors of performance. The following sections explain these concepts

further, with a brief discussion of pitch in music and speech. Due to the concentration of my research on piano performance, pitch is excluded as a variable of investigation. However, although pitch and intensity are not the focus of this study, the complex interaction between the parameters of duration, intensity, and pitch involved in the signalling of structural and performance elements in speech and music means that duration cannot be examined in isolation.

Prominence and Boundaries in Speech

In speech, the acoustic realisation of prominence and boundaries form ‘prosodic cues’. These cues are important because they demarcate cognitive structures within the minds of listeners and speakers that allow for the smoother perception and production of language through speech (Pierrehumbert, 1999). Prosodic structures on a smaller scale (within the word level) are important for lexical understanding, and for the segmentation of words within larger units (Cutler, 1995). Prominence cues at this level vary between languages; in English, word stress, which may (as previously mentioned) denote meaning, is typically conveyed through longer duration, as well as louder and higher (or, at least, simply different, or moving) pitch on emphasised syllables. The same cues can also be used to signal emphasis or focus at the phrase level (e.g. a speaker could lend greater prominence to the final stressed syllable of a phrase by further lengthening that syllable, thus conveying the focus of the phrase). On an even larger scale (that is, above the phrase level), prosodic structures may reflect syntax (principles governing phrase structure in language), semantics (meaning), and discourse structure (Ferreira, 2002). Boundary cues above the word level, such as phrase-final lengthening and pitch declination in English, signify the end of a phrase, and thus a unit of meaning.

Pitch in Speech

Variation of pitch signals a variety of prosodic features, including word stress, lexical tone, or phrase accents in intonational structure. Pitch information can likewise be at the word level (e.g. contributing to prominence cues that signal word, or *lexical*, stress, as in EXport vs. exPORT) as well as above the word level (contributing to intonational phrases that show different information structure; e.g. a declarative utterance and an interrogative utterance will be composed of different strings of pitch accents, or ‘tunes’). As well as grammatical structure, intonation can mark pragmatic as well as emotional meaning in an utterance (Pell, 2001). In some languages, pitch is also used to distinguish lexical meaning. In languages that have lexical tone, such as Mandarin Chinese, tone (word-level semantically contrastive pitch information) interacts with intonation (phrase-level pitch information) in complex ways (Schack, 2000). One theory that may have important implications for my experimental study (presented in Chapter Three) suggests that a tonal language speaker’s perception, and therefore also production, of prosodic cues in an L2 (second or acquired language) non-tonal language are influenced by the tonal characteristics of her or his native language (see Zhang, et al., 2008 for stress perception and Zhang, 2012 for perception of phrase boundaries).

Prominence and Boundaries in Music

Prominence has a number of roles in expressive musical performance: it is employed for melodic, harmonic, or rhythmic emphasis (e.g. by playing certain notes louder and slightly longer than notated, by ‘voicing’ certain notes in a chord by playing them slightly louder

than the others, or slightly exaggerating or changing the articulation, e.g. more *staccato* or *legato*), as well as for showing metrical and phrasal structure. Musical metre is a perceptual phenomenon that involves awareness of a beat or pulse (regularly-occurring articulations in a range from 100ms to 2 seconds apart), as well as a sense of a hierarchy of multiple layers of beats (according to Lerdahl and Jackendoff, 1983, in order for strong or weak beats to be perceived, a metrical hierarchy must exist with at least two or more levels of beats; London, 2001). A 'phrase' is a term that is borrowed from linguistic terminology; it is a musical unit characterised by a melodic line (underpinned by harmony or not), with a sense of beginning and ending. Musical phrases can be subdivided as well as combined to form larger structures, known as periods. The end of a phrase is often signalled by a cadence, which involves melodic and harmonic motion that gives a sense of closure. Performers use prominence cues of timing and intensity to highlight the beginning and ending of a phrase; e.g. by delaying the placement of the first note(s), by an accent (more intensity) or a sudden unaccented note (drop in intensity) in a loud passage. To show metrically important events (such as the 'downbeat' or first beat of a bar), performers may likewise lengthen tones, perform them louder, or change the articulation (Sloboda, 1983; 1986). Prominence is also employed in the emphasising of emotional characteristics within the music, as well as highlighting mood and even function (e.g. 'declamatory' vs. 'questioning'). To highlight certain rhythmic elements in the music, performers may exaggerate them, for example by play durations of a 3:1 ratio (a dotted quaver and a semiquaver) as >3:1 (Gabrielsson, 1987); this emphasis of the length of the first note at the expense of the shorter second note can be for a variety of purposes (to create a feeling of yearning by leaning on the first note, to create a 'swinging' effect for motion, etc.).

In addition to prominence, performers use certain cues to mark the lengths of structural units, which include short melodic or rhythmic figures, phrases, and larger sections of music. Typically, a phrase is heard as a complete musical unit or idea that can then be elaborated, extended, shortened, and transformed in a variety of ways throughout a piece. Phrase boundaries are marked by changes in duration, intensity, and articulation (Henderson, 1936). Several studies have determined that performers use phrase-final lengthening and a gradual softening of dynamic (amplitude) to show the end of a phrase (Gabrielsson, 1987; Kendell and Carterette, 1990; Palmer, 1989; Todd, 1985). This also applies to larger sections of music; slowing down of tempo and decreased amplitude are typically more significant at main cadential points than at phrases within a section (see Shaffer and Todd, 1987; Todd, 1985, 1989 for a hierarchical theory of phrase grouping principles in musical performance).⁸

Pitch in Music

Pitch variation in music can be separated into two categories: composed and expressive. In composed music, pitches are determined by the composer and performers generally do not alter the notated pitches in the score (exceptions occur when improvisation is involved).⁹ Composers may also choose to embed prominence and boundary cues within their scores using pitch; added pitches show prominence (e.g. a full chord rather than a single note; more voices rather than a single line) and lowering pitches at the end of phrases often

⁸ This parallels discourse segmentation of larger sections in speech; see the introduction in Oliveira, Jr (2003).

⁹ Music of the 17th, 18th and early 19th centuries typically featured improvised passages or embellishments; this constitutes a further level of expression that will not be addressed in my work.

denotes boundaries.¹⁰ Expressive pitch variation refers to the slight change in pitch (without altering the pitch category, i.e. the note that is chosen by the composer) made by the performer to emphasise certain elements in the music. For example, slight and rapid variations in pitch (vibrato) can be used to mark a particularly important note in a melody. Also, pitch intervals can be exaggerated in a melody by slightly enlarging the pitch distance between two notes (to show the difference between an augmented fourth and a diminished fifth, for example). But certain musical instruments are extremely limited in their ability to vary (as opposed to change) pitch. The piano is an example; the instrument is unable to vary individual pitches (i.e. make them more 'sharp' or 'flat', depending on the expressive context of the music) or produce pitch vibrato. While the comparison of expressive pitch variation and variation of pitch in speech is of great interest, it is for this reason not applicable to an investigation of prosody in speech and piano performance.¹¹ On the other

¹⁰ In speech, fundamental frequency or F0 (which relates to the rising or lowering of the voice, as determined by vocal fold vibration) produces pitch, while other frequency information (which relates to the configuration of the vocal tract) differentiates between different vowels and also certain consonants; i.e. the difference between 'bat' and 'bet'. Conventionally, frequency information imparts lexical information, while F0 carries prosodic information (whether something is said as a question or statement, or is emphasised in the phrase, etc.). However, paralleling the use of pitch to embed 'musically-prosodic' cues within a composition, speakers can manipulate vowel quality (which is usually constrained because it changes word meaning) to mark prosodic cues. For example, schwas (such as the second syllable in the word, 'water') in English are always unstressed, so when a listener hears a schwa, he or she is getting prosodic information. Similarly, when the function word, 'the' (which is not normally stressed in a phrase and will typically be a schwa) is emphasised, the vowel becomes /i:/, rhyming with 'tea'. This kind of frequency information can be manipulated by the speaker to signal prosody, and is similar to certain expressive pitch decisions made by the composer.

¹¹ That is to say, studies similar to mine but using bowed instruments, wind instruments, or the sung voice

hand, being a pianist allows for particular focus on the variables of duration and intensity, and arguably has the advantage of precluding modulation of pitch as a means of prosodic expression.

2.6 Further Clarification: Speech Rhythm, Musical Rhythm, Musical Prosody, and Nuance

The above considerations of the comparable elements in speech and music have been associated with rhythm; however, it is necessary to note that in musical performance the concept of expressive variation—or micro-deviations from the notated durations, pitches, timbres, and dynamics in the musical score—are part of what is often referred to as ‘nuance’. Such nuance, according to Davidson (2014), can be divided into two contexts: intra-musical and extra-musical. Intra-musical nuance relates not only to structural and rhythmic features but also to the stylistic qualities: different music with the ‘same’ rhythm, for example minim-crotchet figures in Viennese waltzes and Ländler respectively, is executed differently. In addition, various expressive descriptions indicated in the score by the composer (*con anima*, *dolce*, *langsam*, etc.) require deviations of timing relative to the precise, ‘mechanical’ values prescribed by the notation. Extra-musical nuances, on the other hand, involve performer idiosyncrasies as well as physical gestures and movements that may or may not work in tandem with intra-musical decisions. It is possible for performers to develop certain physical and/or musical idiosyncrasies that become tendencies (conscious

would doubtless find pitch an extremely informative (i.e. expressive) variable compared with speech.

or unconscious) or even mannerisms, used for any performance regardless of structure or style.¹²

At this point, it is necessary to summarise the four concepts of speech rhythm, musical rhythm, musical prosody, and nuance in musical performance, as well as to discuss my particular research perspective, which addresses all of these areas and their overlapping elements. According to the view of Post and Payne (2018: 7), speech rhythm is a percept that involves both structural and speaker-based properties that are signalled by multiple parameters (e.g. duration, intensity, and fundamental frequency). Musical rhythm in the tradition of Western classical music commonly refers to the designs of the composer, as well as the specific elements within a composition (e.g. metrical and even supra-metrical organisation as well as small-scale figures and patterns); these compositional decisions have both structural and expressive implications that require a performer to execute and interpret them. Musical prosody, defined as the musically expressive elements of phrasing, *rubato*, and dynamic and articulatory variation in performance, shifts the focus from traditionally score-based conceptualisations of musical rhythm to those exhibited in music as performed and interpreted (Palmer and Hutchins, 2006). In this way, the concept of musical prosody seems to align with Post and Payne's approach to speech rhythm in that musically prosodic elements are likewise 'experienced' as properties that are signalled by acoustical parameters. The concept of nuance, both intra-musical and extra-musical, further focuses upon the performed music and the performer as the executor of expressive variation inherent in the composition, as well as that based on her or his own expressive

¹² Davidson (2014) gives the example of extending the duration of a held note in order to make time for a physical gesture, showing physically as well as musically an intention of expressivity.

tendencies. This expressive variation, according to Palmer and Hutchins's concept of musical prosody, can be both structural and idiosyncratic and may be analysed and investigated according to specific acoustical parameters such as timing (duration), dynamics (intensity), and pitch (fundamental frequency). For my own research, I have decided to focus on the variability of duration.

2.7 Research Review

What follows is a review of studies on expressive musical performance, speech rhythm, and cross-domain commonalities that are particularly relevant to my research. The methodologies and designs found in these studies have been influential in the formation of my own experimental design.

Research on Expressive Timing in Music

A particularly useful measurement in the analysis of timing deviations in human performances is that of 'inter-onset intervals' (IOIs). IOIs are the temporal intervals of successive note onsets and are the basis of 'expressive timing profiles', i.e. series that show the actual pattern of event timing versus the idealised pattern based on notated duration. For example, Repp (1992) studied several pianists' performances of Schumann's *Träumerei*. Approximately 170 quavers were analysed, showing significant variation in IOIs. These showed that all the performances exhibited longer intervals (slowing of tempo) at structural boundaries, with the amount of slowing proportional to the importance of the boundary. More specifically, they revealed that within individual melodic phrases there was a tendency to accelerate at the beginning and slow near the end. Another informative aspect of

expressive timing, although not of focus in my study, is articulation (see Patel, 2008: 115). Articulation relates to the physical duration of a time interval, so that while IOI refers to the time interval between the onsets of successive tones, articulation refers to the time between the offset of one tone and the onset of the next. It is likely for the physical duration of an event to be shorter than its IOI in *staccato*, for instance, and longer in overlapping *legato*.

Several perceptual studies have revealed the following results: both musicians and nonmusicians can reliably identify performances of the same music as ‘expressive’, ‘mechanical’, or ‘exaggerated’ (Kendall and Carterette, 1990) and can identify the performer’s intended emotion based on expressive features (Gabrielsson and Juslin, 1996). Furthermore, it is suggested that musically trained listeners can identify the intended metrical and phrase structure of a performance based on expressive cues (see Palmer, 1996). Clarke (1993) extracted the expressive timing profiles of ‘naturally’ performed short melodies, manipulated them, and then re-imposed them onto a mechanical performance of the melody, thus creating versions of the melody with mismatched structure and expression. Musicians judged the originals versus the mismatched melodies in terms of the quality of the performance, and favoured the originals. This confirms that informed listeners are sensitive to the way expressive timing aligns with the structure of musical passages.

Just as with earlier linguistic theories of normalisation, which regarded prosodic variation as ‘noise’ in the speech percept that is ‘normalised away’ (see Pisoni, 1997 for an overview), it had been suggested that individualistic aspects of music performance, such as expressive timing patterns, are ‘reduced’ away in listeners’ memories of musical sequences (Large et al, 1995). Thus, it was thought that abstract memory representation favours a less detailed, more categorical structure. More recent research suggests, however, that listeners

retain certain temporal information in memory for both speech and music (Bradlow, et al., 1999; Palmer et al., 2001). In Palmer's study, listeners were familiarised with performances of short melodic sequences, and then tested for their ability to recognise these performances compared with others of the same sequences. A pianist performed the different performances of the same short melodic sequences as part of longer melodies with differing metrical structure; because of this, the performances exhibited different timing as well as articulation and intensity patterns. Results suggested that both musicians and nonmusicians could recognise the original version compared with others.

Research on Speech Rhythm

Speech rhythm typology is a research approach that seeks to understand the similarities and differences among the world's languages. Impressionistically, languages may be said to fall within certain rhythmic groups; speech rhythm typology investigates this idea through four approaches: isochronous, phonological, durational, and perceptual (see Patel, 2008 for a summary of typological research). The first approach, of isochrony or periodicity, has long been disproven, but it has introduced terminology that is still in use, albeit typically only as shorthand and not in a literal sense. The categorical terms 'stress-timed' and 'syllable-timed' were coined by Kenneth Pike (1945), who proposed that languages could be divided into two rhythmic groups depending on the patterns of syllables and stress. Syllable-timed languages (such as Spanish) were said to exhibit equal timing between syllables, while stress-timed languages (like English) had equal timing between stresses. Despite variation of the number of syllables between each stress, speakers of stress-timed languages would elongate or shorten syllables to fit into a pattern of evenly-spaced intervals between stresses (called 'feet'). Abercrombie (1967) developed Pike's contention by explicitly stating

that every language in the world is spoken with some kind of rhythm, and that the rhythms recur periodically. This idea resembles the 'beat' concept in music; Abercrombie seemed to be asserting that languages have rhythmic pulses just as music does.

Empirical research on isochrony of stresses or syllables in languages, however, failed to produce any evidence in the acoustic signal. In one study (Roach, 1982), English and Russian could not be discriminated from syllable-timed languages such as French and Telugu on the basis of inter-stress timings. In another (Lehiste, 1977), English stress feet were actually found to grow in duration as syllables increased, contrary to Pike's theory. Patel (2008: 121) suggests that the reason why Pike's terminology for language groups still exists is that it somehow 'matches subjective intuitions about rhythm'. Likewise, Beckman (1992) believes that certain languages are perceived as rhythmically similar, even if the basis for that perception has not been understood.

Although the search for isochrony seems to have failed, the percept appears strong enough to persuade researchers to seek other typological explanations. The phonological approach views languages within different rhythmic groups as having phonologically different properties that influence how they are organised. In research that came to be central to the phonological approach, Dauer (1983, 1987) suggested that stress-timed and syllable-timed languages differ in their phonological properties such as the diversity of syllable structures in speech, the existence of vowel reduction, and the influence of stress on vowel reduction. English, for example, has been suggested to have more variable syllable durations than French (a so-called syllable-timed language). Furthermore, vowel reduction exists in English as opposed to Spanish (aforementioned as syllable-timed), and vowels in stressed syllables of English become longer than the same vowels in unstressed syllables; in Spanish, stress does not affect vowel duration to such an extent. Bolinger (1981) theorised

upon the temporal patterning in languages (focusing on English) from another perspective. He believed that the rhythm of English speech is characterised by an alternating pattern of syllables with full and reduced vowels. His work also suggested that English has two kinds of rhythm structured as two levels of temporal patterns; in other words, above the level of alternating long and short syllables exists a second level of rhythmic patterning based on pitch accents. This two-levelled approach to rhythmic structure resembles the concept of musical metre.

Based on (and moving beyond) the work of Dauer and Bolinger, Ramus and colleagues (1999) and Low, et al. (2000) focused on examining the durations of vowels, consonants, and their patterns in speech. In particular, they tried to quantify the variability in these durations. Ramus et al. (1999) developed 'rhythm metrics' that expressed the extent of 'vocalicness' (percentage of vowel duration, or %V), as well as the variability of consonantal and vocalic intervals (ΔC and ΔV) within sentences for different languages. Low, et al. (2000), inspired by Bolinger's idea of vowel reduction and alternation, developed the raw and normalised 'pairwise variability index' (rPVI and nPVI), which calculates the difference in duration between each pair of successive measurements. The measurements in use are either vocalic or consonantal interval durations (n.b. an interval may have more than one vowel or consonant) within an utterance, while normalising the data means taking the absolute value of the difference and dividing it by the mean duration of the pair (Grabe and Low, 2002), which eliminates variables such as speech rate among speakers. Grabe and Low used the nPVI to examine vowel durations in sentences of several languages and showed that 'stress-timed' languages (such as German, English, and Thai) had higher nPVI than languages classified as 'syllable-timed' (French, Italian, Spanish). This process was then

used by Patel and Daniele (2003) to extract empirical data from musical themes, leading to significant implications for the music/language relationship (more on this in Chapter Three).

The Rhythm Class Hypothesis (which proposes a categorical distinction between the rhythms of certain languages) and associated rhythm metrics (see White and Mattys, 2007 for a comprehensive overview) have been strongly critiqued following the observation that different studies yield different metric scores and can sometimes place a given language in a different class. The metrics vary for different tasks, speakers, contexts, and experiments (see the work of Wiget et al., 2010; Grabe and Low, 2002; White et al., 2009; Arvaniti, 2009, 2012), and many have argued that the metrics are not robust and therefore not useful for speech rhythm measurement. However, the issue seems to stem from the failure of the Rhythm Class Hypothesis, and not of the metrics themselves. As Post and Payne (2018: 6) argue, the metrics should be seen as a separate and ‘neutral tool for comparing holistic timing characteristics, many or all of which contribute to linguistic discrimination’. The metrics may reflect any and all sources of durational variability in speech, including those that are phonological, prosodic, and also performance- and speaker-specific. Thus, while metrics will never be identically robust across contexts, they are a useful comparative measure. Duration, intensity, and pitch (among other elements) seem inextricably connected in both speech and music percepts, and while metrics are only capable of capturing temporal properties, these durational measurements are potentially influenced by multiple systems. This makes metrics a promising tool for comparing acoustic variability within and across speech and musical performance.

As previously mentioned, certain metrics have caught the attention of cross-domain researchers such as Patel and Daniele (2003) and, subsequently, Huron and Ollen (2003). The nPVI is cited as useful for two reasons (Patel, 2008): one, it measures *relative* durational

contrast; i.e. 'the durational difference between each pair of intervals is measured relative to the average duration of the pair', which can control for differences in both speech rates and musical tempi; and two, it has been applied to vowels in speech (the core of syllables), which can be compared with musical tones. The following section dives further into cross-domain, as well as cross-cultural research, in which the work of Patel and others are located.

Cross-Domain, Cross-Cultural Research

Researchers have begun to examine the relationship between speech and music across language and culture, particularly within perception. Of the studies on cross-linguistic comparisons between musicians and non-musicians (see Sadakata and Sekiyama, 2011 for a summary), it has been shown that musically-trained participants out-perform non-musicians in tasks involving verbal memory (Ho et al., 2003; Jakobson et al., 2008), speech production (Slevc and Miyake, 2006), and reading (Anvari et al., 2002). Significantly, studies have shown that musical training has the most effect on the perception of both linguistic and non-linguistic timing (Yee et al., 1994; Rammsayer and Altenmüller, 2006). Speech timing studies include perception of vowel durations (Milovanov et al., 2009), metric structure (Marie et al., 2011), and discrimination/identification tasks using sounds from different languages (in this case, Japanese and Dutch; Sadakata and Sekiyama, 2011). According to Sadakata and Sekiyama (2011: 8), musicians' enhanced ability to perceive timing in speech may have to do with the perception of 'equally prominent' elements of both language and music, such as duration and pitch perception, that is developed during musical training.

An alternative line of research involves the investigation of music written by composers of different cultures, examining the influence of speech rhythms upon musical

themes. Patel (2008) expresses an idea that has been of interest to linguists, music scholars, as well as performers throughout history (see Hall, 1953; Abraham, 1974; Wenk, 1987; and Garfias, 1987):

It is known from studies of language acquisition that the perceptual system is sensitive to the rhythmic patterns of language from a very early age (Nazzi et al., 1998; Ramus, 2002). Composers, like other members of their culture, internalize these patterns as part of learning to speak their native language [...] We suggest that when composers write music, linguistic rhythms are ‘in their ears’, and they can consciously or unconsciously draw on these patterns in weaving the sonic fabric of their music (Patel, 2008: 165).

Examination of speech rhythms in music began with Wenk (1987), who was influenced by the writings of English musicologist Gerald Abraham. Abraham (1974: 18) dedicated an entire chapter in his book to the ‘factors of language’; his anecdotal considerations of the influence of a nation’s language upon its music (regarding Italian, Czech, Russian, French, German, and English music and composers) inspired Wenk to take a more methodological, linguistic approach. Focusing on French (‘syllable-timed’) and English (‘stress-timed’) language, Wenk noticed that French speech shows a tendency for final syllable lengthening. He characterised this lengthening as a type of accent (with duration but not intensity) that is perceived at the end of a rhythmic group. On the other hand, ‘stressed’ syllables in English are perceived at the beginning of their rhythmic group. An example of this is given in Wenk (1982: 192):

a. *Phillipe étudie à l’université*

b. *Phillip’s studying at the university*

(accented syllables in italics)

Wenk predicted that French music would have a higher propensity for phrase-final lengthening than English music. He had a professional musician mark the phrase boundaries in English and French classical music and found that there were a larger number of phrases in the French music in which the final note was the longest note in the phrase. The drawbacks of this study were a lack of empirical evidence to support his findings (a professional musician was used to mark French and English musical phrases), a limited sample size (only one composer was used from each language group—Poulenc and Britten, and only one movement from one piece was used from each composer), and the decision to use vocal music set to Latin texts instead of music involving the native language in question. Wenk's study, though seemingly flawed, paved way for future empirical inquiries of a similar nature. Patel and colleagues (see Patel and Daniele, 2003; Huron and Ollen, 2003; Patel et al., 2006) measured numerous themes from a standard musicological reference (Barlow and Morgenstern, 1983) by assigning a duration of 1 to the first note of each theme, with the durations of the remaining notes expressed as a fraction or a multiple of 1. These durations were analysed using rhythmic metrics to compare for variability of lengths between consecutive pairs of notes. It was determined that music with higher nPVI such as American, English, and Swedish came from cultures with high linguistic nPVI whereas French, Italian, and Spanish music had significantly lower nPVI, corresponding to their languages. Additionally, speech melody was tested for its influence on French and English musical melodies (Patel et al., 2006). It was found that pitch interval size variability in speech was also reflected in music: French speech has lower pitch interval variability than English speech and the same was revealed of the intervals in music. This purports to confirm the tendency for French speech and music to have smaller intervals between successive pitches

and tones, respectively, compared with English speech and music. At this point, it is interesting to mention observations made by Abraham on French music based on its language and 'temperament': 'a fairly quick patter of even notes of short value, small intervals and narrow tessitura, absence of metrical stresses, everything very supple and subtle' (Abraham, 1974: 256). While these seemingly superficial impressions may be made by anyone with enough familiarity with French repertoire, one cannot help relating them to similar observations made as a result of empirical study; the 'even notes' observation brings to mind the lower nPVI of French music and the 'small intervals' align with the lower interval size variability between consecutive tones. The interpretation of a nation's music based on the prosodic characteristics of its language is certainly a compelling theory, but also problematic: a nation can of course have more than one language, and a language can be of more than one nation. Even where one language is spoken, there will be dialectal variation that may be prosodically quite distinct. Also, it is difficult to determine how much linguistic patterns influence a composer's work, as well as separating this from other factors. The influence of one's own language background upon musical performance therefore seems more fertile terrain for exploration.

Perception of rhythm involves certain principles shared among speakers of all languages; one such principle is that sounds are grouped by the listener to form higher-level patterns. Streams of sound in speech and music are processed by the listener as chunks of information such as words, phrases, and motives, which then allow the listener to recognise bigger patterns and structures. A significant theory first proposed by Bolton (1894) and supported by Woodrow (1909) claims that listeners tend to group sounds varying in intensity into strong-weak patterns while sounds varying in duration tend to be grouped into short-long patterns. This theory has been widely supported by research in sound and

speech perception, and similar principles have been found in music. Recently, however, the validity of this theory cross-culturally has been the subject of debate: Kusumoto and Moreton (1997) and Iversen et al., (2008) investigated the perception of rhythmic groupings among Japanese and English speakers, producing results that contradicted previous findings: the Japanese listeners perceived patterns of alternating duration *differently* from the English listeners; the English listeners generally had a strong preference for short-long grouping (with only a single listener preferring long-short) while Japanese listeners showed more preference for long-short (45%) rather than short-long (26%). On the other hand, a recent study by Langus et al., (2016) suggests that while listeners' native languages influence their grouping of linguistic stimuli, there was no evidence of influence on non-linguistic grouping of sine-wave speech sounds and visual events (participants of this study were speakers of Italian, Turkish, and Persian). Debate on the nature of the influence of speech rhythm on non-linguistic rhythm perception certainly does not preclude the possibility of cross-domain relationships; further studies of this nature are anticipated.

Finding Evidence for Speech Influence in Musical Performance: A Critique of Current Studies

There have often been (largely impressionistic and anecdotal) observations that classical musicians express music differently depending on their culture. One instance from piano pedagogy comes to mind: Theodor Leschetizky and his observations on the varying characteristics of piano students depending on their nationalities. According to Annette Hullah, one of Leschetizky's teaching assistants, the Polish pianist and teacher catalogued all of his students and knew what 'gifts' they would have based on country of origin:

From the English he expects good musicians, good workers, and bad executants; doing by work what the Slav does by instinct; their heads serving them better than their hearts. The Americans he finds more spontaneous. Accustomed to keep all their faculties in readiness for the unexpected, their perceptions are quick, and they possess considerable technical facility... The Russians stand first in Leschetisky's opinion. United to a prodigious technique, they have passion, dramatic power, elemental force, and extraordinary vitality... The Pole, less strong and rugged than the Russian, leans more to the poetical side of music. Originality is to be found in all he does; refinement, an exquisite tenderness, and instinctive rhythm. The French he compares to birds of passage, flying lightly up in the clouds, unconscious of what lies below. They are dainty, crisp, clear-cut in their playing, and they phrase well. The Germans he respects for their earnestness, their patient devotion to detail, their orderliness, and intense and humble love of their art. But their outlook is a little grey. The gentle Swedes, in whom he finds much talent, are more sympathetic to him; and the Italian he loves, because he is Italian—though he cannot, as a rule, play the piano in the very least (Hullah, 1906: 72).

While such stereotypical observations refer more to cultural character and are especially outdated in today's politically correct social climate, artefacts of such categorisation still remain within the minds of teachers today, and are passed down to their

students.¹³ From there, it is but a short step into linguistic territory due to the significance of language within any culture: could the observations of a highly influential piano pedagogue on national character also be related to linguistic influences on the expressive qualities of his students? Sadakata et al. (2004) investigated the subjective observation that Japanese and Western musicians perform Western classical music differently,¹⁴ citing previous production studies involving the rhythmic analysis of pianists' performances of Mozart's Piano Sonata in A major, K 331. In one of these, Gabrielsson (1987) analysed the timings of five pianists' recordings and noticed that all tended to shorten the semiquaver note and elongate the surrounding dotted quaver and quaver in the rhythm , except for one Japanese pianist who sometimes elongated the semiquaver and shortened the surrounding notes. Focusing on this distinction, Ohgushi (1999, 2002) compared the intervallic durations of Japanese and Western pianists playing the same Mozart sonata. He discovered that there was a discrepancy between the two groups concerning the performance of three intervals with a rhythmic ratio of (3:1:2). Similar to the data from Gabrielsson, Ohgushi found that Japanese pianists played the first two intervals (3:1) with a

¹³ I have been told during piano lessons about the particular way that Polish pianists execute melodic content; care is put into every note in a way that is highly distinctive. Hungarian pianists are known for playing with rhythmic vitality and East Asian pianists are not likely to be impulsive and/or courageously expressive due to such behaviour having been discouraged during their upbringing. Of course, it is understood by both teacher and pupil that these are gross generalisations, yet they continue to be discussed.

¹⁴ According to Sadakata et al., (2004: 390), Hideo Saito, a Japanese conductor, described a 'specifically Japanese interpretation' of Mozart's pieces taught by Japanese college-level teachers which seems to reflect the Japanese language. Minao Shibata, a composer, recalled how young Japanese musicians felt that the sensitivity of Japanese musicians was different from those of Westerners.

shorter intervallic duration (in other words, the long note was slightly shorter) than Western pianists.

Sadakata et al. (2004) investigated Japanese and Dutch percussionists trained in the Western classical tradition. Here they sought a more rigorous experimental procedure: the percussionists were asked to play rhythmic ratios of 1:1, 1:2, 1:3, 1:4, 1:5, as well as reverse ratios in two modes, 'mechanical' and 'musical'. In fact, the modes had little effect on the results of the performed durations; this may be due to a lack of musical context in order to stimulate the musicians to make expressive choices (the difference between playing an excerpt from an established piece within the Western classical repertoire such as the Mozart sonata and four bars of an unchanging rhythmic pattern is obvious). Results for the expected effect of cultural background upon certain ratios were also less conclusive. The Japanese percussionists were expected to play the 3:1 pattern with a smaller ratio than Dutch percussionists, as implied in the studies by Gabrielsson and Ohgushi. Results showed that although there were significant differences between participant groups in some ratios, they were not systematic. The average duration ratio of Japanese performances of 3:1 was actually larger than that of the Dutch, although the Japanese group tended to play smaller than indicated ratios of 4:1, 5:1, 1:4, and 1:5. This implies that cultural differences in performance timing may be likely to occur in more complex rhythms requiring more awkward (and larger numbers of) subdivisions.

The above study and its results are inconclusive for several reasons: first, the two participant groups differed in the amount of musical training experience and professionalisation. The Japanese group consisted of five undergraduate and one graduate-level percussion majors at an arts university, whereas the Dutch group consisted entirely of professional musicians. The Japanese group averaged 17.8 years of musical training while

the Dutch group averaged 21.5 years. While the difference of averaged years of training is not significantly large, the difference in the degree of professionalisation could have had an impact upon performance skill during the study. Further factors include differing musical obligations and environmental exposure for students compared with professionals. Second, as mentioned above, the use of rhythmic patterns out of musical context is highly problematic for investigating durational variation. It is very difficult for a trained musician to produce a 'musical' performance of a series of rhythms without the cues of phrasing, harmony, dynamic change, articulation, cadential points, and association with style and/or emotion. Classical percussionists achieve expression in performance by modifying timing, articulation, stroke intensities, stick heights, and strike positions (Prockup, et al., 2013). Such a wide range of expressive techniques cannot arise from rhythms which are by themselves inherently unmusical. Also problematic is the use of complex rhythmic patterns foreign to Western classical music (such as the ratios of 1:4, 4:1, 1:5, and 5:1). Although an interesting tendency emerged from the Japanese group, it is difficult to conclude whether and how such variations would appear in performances of standard Western classical repertoire. Finally, although the statement from Hideo Saito (see footnote 10) referred to a possible influence of the Japanese language within the interpretations of Japanese musicians, Sadakata et al. (2004) did not investigate the speech of the participants, thus leaving questions open about what the supposed 'cultural differences' between Japanese and Western musicians may be.

A later study conducted by Slobodian (2008) addressed several questions pertinent to my own inquiry, including the possibility of speech rhythm influence on perception and production of musical rhythm. Korean and English middle school students were tested for their Korean and English speech, their ability to reproduce rhythms of 2:1 and 3:1 ratio by

clapping, and their ability to clap along to the beat of various rhythms. The results of the study were inconclusive, with Korean and English participants having little rhythmic difference in speech of Korean, English, and Emakuah (control language) across language groups as well as in rhythm production and perception. This is most likely due to several methodological issues and variables: one, each experiment with a participant lasted no longer than seven minutes in order to minimise missed class time; two, the use of clapping rather than tapping pads reduced accuracy for data analysis; three, one of the tasks (clapping to the strong beat of rhythms) was confusing for both Korean and English students and resulted in a wide range of results with some clapping to every beat, some clapping to the exact rhythm heard, and some not clapping at all; and four, the participants varied in the range of musical experience and exposure to Western classical and Korean traditional music. In a review of Slobodian's study, Iverson (2008) points out the importance, when dealing with multiculturalism in rhythm research, of choosing cultures with contrastive languages (syllable-timed and stress-timed or tonal and non-tonal, for instance); of articulating the characteristics of each language; and of identifying the degree of each participant's musical training or exposure to music. The latter may have a significant effect on rhythm perception and production, and hence on task performance success and quality.

Returning to the study by Sadakata et al. (2004), a significant concern relates to its lack of ecological validity. Musicians within a performance context function as specialised individuals whose behaviour is determined by experience, training, and the environment around them. Thus, research on the performance of musical rhythm benefits from the use of real-time performance situations and samples rather than isolated rhythmic events that lack musical context. Metre, for example, has significant influence upon the execution of relative durations; musical style and character are also factors. In the case of Slobodian's

study, the significantly large sample size (134 middle-school students) proved to be a challenging factor: students were varied in their musical experience and exposure, which resulted in varied quality of task performance. Furthermore, as the students were mostly concentrated on general education or non-musical extra-curricular activities, participant time had to be as short as possible in order to minimise absence from regularly scheduled class time. These challenges further underline the necessity, when testing for musical production, of controlling for musical experience and ability. Also, narrowing the sample to involve only participants with music as a significant or primary focus will generally allow for longer and more in-depth observational periods.

In the next chapter, I present my own experimental approach to investigating cross-cultural and cross-domain factors, with a goal similar to those of the above studies: to investigate influence of language-specific speech factors on musical performance. However, my design focuses on comparing, with equal rigour, the speech (Mandarin Chinese and English) of classically trained pianists and their musical performances. To my knowledge, comparative studies of this specificity are rare.¹⁵

¹⁵ For a cross-cultural study on the *perception* of Chinese and German classical musicians, see Nan, et al. (2006). Results of this study suggest that, due to the influence of acculturation, Chinese and German musicians were equally able to perceive and identify Western classical music.

Chapter Three: Experiment

3.1 Introduction

This chapter describes the empirical experimentation that I undertook. Motivated largely by the lack of empirical interdisciplinary literature that could address my particular research questions, I adopted a quantitative approach. I was also driven by a curiosity (and a rather ambitious aspiration) to see how far I could push the boundaries of my own background in classical musical performance by utilising methods taken from a very different discipline: phonetic science. This chapter follows a format typical of research using experimental methodologies, adhering to the standards of the scientific process as far as possible. My ultimate goal, in addition to obtaining a set of quantitative results to support my hypotheses, was to explore the contribution that an analytical, empirical study can make to broader research questions about expressive musical performance, and to make this analytical part of my study methodologically transparent and replicable. I hope that this may also motivate those in the disciplines of both music and linguistics to conduct further interdisciplinary research.

The primary purpose of this study has been to investigate whether empirical data may be produced in support of the hypothesis of linguistic influence upon musical performance. Specifically, I was interested in whether a pianist's speech prosody, which may to some extent be determined by the prosodic structure of their native language, influences, or is in any way correlated with, the musical expression of her or his performance. My specific research questions are as follows:

- i) What are the comparable, measureable variables involved in speech prosody and musical expression in performance?

- ii) Are there significant systematic differences in these variables that would suggest that native language speech prosody behaviour may be reflected in expressive musical performance?

To address the first question, I conducted a pilot study with six conservatoire-trained classical pianists to determine the comparable variables involved in both speech prosody and musical expression. The pianists were divided into two native language groups: Standard Mandarin Chinese and English. Based on the results of the pilot study, and in addition to the review of relevant literature (mentioned in the previous chapter), I narrowed my focus to durational variability, which is a central component of the signalling of prosodic cues in speech prosody and expressive cues in musical performance.

As for the second question, I hypothesised that I would find systematic differences in the degree of musical timing variability (that is, above and beyond what is prescribed by the composed notation) between language groups. Specifically, I expected those in the native English group to perform with a higher degree of durational variability *between successive quaver beat durations* than those in the Mandarin group, as a result of the influence of structural timing patterns in their respective native languages. Alternatively, I hypothesised that there may be an effect at the level of the individual, manifested through consistency between the timing variability of both the speech and musical performance *within speakers'* individual data, which would suggest that expression is an individual, performative capacity that affects both speech and music. These two hypotheses were formed under the assumption that durational variability in speech and music are not unrelated, and this is either because a) expressive durational variability in musical performance is influenced by durational variability in the background language, revealing a language-effect, or b) durational variability in both musical and linguistic performance are underpinned by unitary,

general performance ability, supporting an individual-effect. A third and final hypothesis was that there would be no correlation between durational variability in musical performance and language group, suggesting that musical expressivity is an individual capacity that is unrelated to language background. The results of my study seem to disprove the third hypothesis while suggesting an additive possibility of the first two: that durational variability in music and speech are independent phenomena that may interact *both* within speakers and between language groups.

This chapter outlines my experimental methodology. In the following sections, I explain my experimental design while considering my particular role as both a ‘practitioner’—and ‘insider’—researcher; I also elucidate my pilot study and the variables determined from this process; additionally, I discuss ethical considerations and certain limitations of this research. Then follow the results and analyses of my experimental data. The concluding section refers to the multi-componential model of speech rhythm, outlined by Post and Payne (2018), in order to ground in current theory a possible explanation, or at least a framework, for the experimental results. This leads to the turning point that is Chapter Four, where I analyse and interpret the results of my experiment from a performer’s perspective by use of autoethnography.

3.2 Experimental Design

Background and Overview

The idea that the prosody of a composer’s native language can influence the structure of her or his instrumental music has long been suggested by both musicologists and linguists (e.g. Abraham, 1974; Heffner, 1950; Wenk, 1987). It had proved difficult, however, to move

beyond impressionistic statements; empirical data were lacking, and investigations on the influence of language upon composed, instrumental music largely remained within research frameworks founded in qualitative, rather than quantitative methods. One of the reasons for this separation involves the difficulty of developing and applying comparable, quantitative measures to elements found in both music and speech, such as melody and rhythm.¹⁶ As mentioned in the previous chapter, Patel and Daniele (2003) developed a comparative, empirical method to test the claim that native speech prosody can influence the structure of *composed*, instrumental music, with suggestive results. They state three requirements for such a method:

- i) A measure, able to quantify prosodic structure in one or more languages, must be used.
- ii) This same measure must be applicable to music so that music and speech can be compared in a common framework.
- iii) Samples of both music and speech must be broad enough to ensure that results are not idiosyncratic to particular speakers and/or composers.

Patel and Daniele focused their attention on rhythm, or rather regularities and irregularities in timing relations, and were particularly interested in the metric developed by Low et al. (2000) that quantifies the durational variability between durations of successive vocalic intervals of speech for a particular language. This metric, nPVI, is defined as

¹⁶ This itself is also in part due to typical discrepancies in definition and in the use of labelling of these concepts.

$$\text{nPVI} = \frac{100}{m-1} \times \sum_{k=1}^{m-1} \left| \frac{d_k - d_{k+1}}{\frac{d_k + d_{k+1}}{2}} \right|$$

.¹⁷ It satisfies the first two requirements stated above in the following ways: one, nPVI is claimed to be an effective measure for the percept of speech rhythm (or at least of evenness of timing) because it is a *relative* measure of variability; i.e. the durational difference between each pair of intervals is measured relative to the duration of the pair, and this controls for fluctuations in speech rate; and two, nPVI can be applied to vocalic intervals (which cover the core of syllables, agreed to be most sensitive to prosodic effects in speech), but can fairly straightforwardly also be applied to tone intervals (defined as any steady, measureable sound; in this case, musical notes).¹⁸ As for the third requirement, Patel and Daniele (2003) used previously-obtained representative nPVI values for English and French utterances from Ramus (2000)'s study.¹⁹ For the music part, they sourced a total of 137 English instrumental themes and 181 French instrumental themes (composed by six English and 10 French composers) from *A Dictionary of Musical Themes, 2nd Edition* (Barlow and Morgenstern, 1983). Following Patel and Daniele's study, Huron and Ollen (2003) expanded upon the musical samples to include 2000 themes. The findings of these studies, and subsequent studies following their methods, provide evidence to suggest that the nPVI

¹⁷ This version followed their initial development of the raw Pairwise Variability Index (rPVI), which did not account for speech rate.

¹⁸ Patel and Daniele (2003) reason that in the setting of words to music, it is common for one musical note to be assigned one syllable.

¹⁹ At the time of the cited article, this was the largest published set of nPVI values for British English and French.

values of the themes of composers' instrumental music are comparable to the nPVI values of their respective native languages.

I have given an overview of the work of Patel et al. because it is among the first to take a quantified comparative approach, motivating scholars to continue down the path of comparative empirical studies (e.g. Raju, Asu, and Ross, 2010; McGowan and Levitt, 2011; London and Jones, 2011; Carpenter and Levitt, 2016). However, one of the limits of this work is its use of musical texts (that is, measuring the duration of note values prescribed by the composer in a score) along with recorded speech. Since timing in music may be notated, while in speech it typically cannot, and since speech is based on performance (and only in part dependent on the prescribed norms, i.e. prosodic rules, of the language), while musical notation is based on the absolute prescriptions of the composition (that await realisation through performance), this makes for a problematic comparison. So, while I have been influenced by this work, I have taken further steps towards creating a method that looks at the performance of both language and music, and thus makes for what I argue to be a more valid comparison. This method involves data extraction from both recorded musical performance and read (i.e. performed) speech, and seeks to treat the disciplines of music performance studies and speech studies with equal focus. As such, my experimental design uses established speech-measuring tools and processes, as well as technology for the analysis of recorded music, developed within the field of musicology. The open-source application *Sonic Visualiser* (Cannam et al., 2010), allows for the visualisation, annotation, and analysis of recorded music files, much like software such as *Praat* does for speech. Researchers are thus able to quantify elements within expressive musical performance that have previously been difficult to measure, such as *rubato* (tempo variation) and dynamics (intensity variation). This is particularly useful for historical studies of Western classical

music performance throughout the age of recordings (see Leech-Wilkinson, 2009), as well as comparative studies of different performances of a single piece of music (see Cook, 2007 and Leech-Wilkinson, 2007). Regarding my own experiment, this means that, rather than extracting nPVI measures from score-based musical notation, as others have done (e.g. Temperley and Temperley, 2011), I have been able to calculate nPVI measures of performed music, and therefore investigate not just the supposedly inherent durational variability of a given composed piece, but—critically—also investigate the durational variability arising from an individual’s expressive interpretation. In addition to nPVI, I have also used a rhythmic measure developed by Dellwo (2006) that determines the standard deviation of vocalic interval duration throughout an entire utterance (or in my case, an excerpt of performed music); that is, as a measure of global variability, not localised to successive syllables. This measure, VarcoV, is normalised to control for rate of performance, like the nPVI.

Furthermore, my work differs from previous scholars who have also used performed music (e.g. Raju et al., 2010; McGowan and Levitt, 2011; Carpenter and Levitt, 2016) in that, while they have measured the durations of *each note* of their musical samples, I focused on the durations of each *quaver beat*. In other words, I chose to examine *rubato*, i.e. tempo fluctuations present in expressive performance by measuring durational variability between intervals of quaver beat durations, rather than between notes. This is significant because, just as there may be more than one vowel or consonant within an interval of a vocalic or consonantal duration, there may be multiple notes within an interval of a beat (the larger the beat, the more notes are grouped within an interval); my interest was in those intervals, rather than the specific constituents within them. The purpose of my focus upon intervals was two-fold. First, as I am using the same excerpt of music, compared between

performances, I required a measure sensitive enough to subtle changes of duration above and beyond those determined by the composer through notation, but also practical enough to measure within the time constraints of my study. Furthermore, visual comparisons of the timing contours of the performances would be more difficult to distinguish from one another if note durations were measured, as the visuals would reflect (in overwhelming detail) the inherent contour determined by the prescribed notation, rather than an outline of the expressive timing variation within the performance. By measuring beat durations, I can still account for timing differences between notes without having to measure each note, as the duration of each beat would reflect the lengthening or shortening of notes within the beat.

Second, I wanted to challenge the third stipulation made by Patel and Daniele (2003), which seeks to ensure that resulting data do not represent incidental within-subject phenomena particular to individual speakers and composers. As my research is concerned specifically with the subtle, 'idiosyncratic' phenomena that occurs *between* performers of the same musical work, as well as 'speech performers' of the same textual words, it was necessary for me to develop both a measure and a method that can detect and display such information. Essentially, I have altered the research question of Patel et al. to the following: 'Can the prosody of a *performer's* native language influence the *expression* of her or his *musical performance*?'.

To summarise briefly, my experimental research involved the speech and performance of eleven conservatoire-trained classical pianists, divided into two native language groups: Standard Mandarin Chinese and English. My methodology measured two sets of data: variability in vocalic and consonantal durations in speech and quaver beat durations in

musical performance. These data sets were then subjected to analysis through the calculation of rhythm metrics, including nPVI and Varco measures.

Practitioner and Insider Researcher

Before describing the specific elements of my study, I must also consider the impact of my position as both a ‘practitioner-researcher’ and an ‘insider-researcher’. I will reserve discussion of the position as it pertains to a qualitative approach for Chapter 4; in this more scientifically guided part of the research process, it has influenced my decisions mainly in respect of experimental design, namely my choice of materials, participants, instrumentation, and empirical setting. It is my contention that my position as both a practitioner and an ‘insider’ gives my work an advantage over other cross-domain, cross-linguistic research (see my critique of existing studies in Chapter Two): my ‘practitioner’s knowledge’ within the field of classical music performance provides more ecological validity to my experiment through, among other things, my ability to establish a ‘musician-friendly’ setting and process of elicitation. My participants are performing stylistically familiar music on a good-quality grand piano in their own conservatoire, and in the presence of a classmate and fellow performing musician. Furthermore, my analysis of the performances and discussion of the experimental results are informed by my own expert musical knowledge and sensitivities, whereas other studies have frequently needed the validation of an external musical advisor.

To spell out the details of my insider status: during the time of the pilot, as well as the subsequently adjusted main experiment, I was (and still am now) an active practitioner of Western classical music; I practiced the piano, engaged with teachers and peers, and performed in concerts, lessons, and masterclasses throughout my research. Furthermore,

throughout this process, I was a student at the institution in which I conducted my experiments. My personal background, as introduced in Chapter One, involves experience in both Mandarin Chinese and English speech; Mandarin Chinese is my first language and English is currently my predominant language. All of this implies that, by conducting research on pianists' speech and musical performance within my own institutional setting, I am functioning as a practitioner- and insider-researcher, with intimate knowledge and practical experience of both my empirical setting and empirical field. With this unique position come certain advantages and disadvantages.

In quantitative research, it is often assumed that the researcher should maintain a detached, 'objective' relation to the object or objects being observed. This is informed by Positivist philosophy, associated with the gathering of 'factual' knowledge in the form of collected and measurable data. Additionally, a positivist approach requires, to an extent, that the researcher maintain minimal interaction with the research participants, and remain detached and neutral about what is researched. Similarly, an 'outsider' conducting even qualitative research is usually assumed to be able to keep sufficient distance from what is observed to be able to deliver a demonstrably objective account of it. My position as both a practitioner and an insider prevents me from satisfying these positivistic requirements,²⁰ yet I believe that my very particular position not only carries more advantages than disadvantages, but is in fact necessary to enter fully into the complexity of the data generated by my research questions.

²⁰ It must be stated that the existence of a truly 'objective' position within research, regardless of method, is dubious; the researcher, regardless of experience and/or relationship with the experimental field and setting, views the world with shifting and acknowledged experiences, assumptions, and knowledge (Ford, 2010: 72).

To begin with the disadvantages: my position as an ‘insider’ means that I have a certain relationship with the participants of my study. As the participants are current or former students of my conservatoire, and are therefore my peers and acquaintances, I inevitably have existing knowledge of their playing and, on a superficial level, their musical abilities. To control for this, I refrained from making any judgments of the performances from a musical perspective during the data collection process, acting as a neutral audience during the participants’ performances and only making comments when errors in the realisation of the notation were made.²¹ As my experiment is measured quantitatively, I could rely on the resulting data to suggest interpretations at this stage. Another concern that results from my ‘insider’ position relates to the behaviour of the participants towards my presence during the recording sessions. Rather than being in this sense an ‘objective’ researcher, or one with little or no professional training in music, or even an external researcher from another musical institution, I have both been trained within and am intimately familiar with the conservatoire in question. Some may wonder if this could have affected the results, perhaps if the participants misconstrued my presence as implicit or potential judgment of their musical abilities.

My defense against this comes from the very fact that I am an ‘insider’; I am aware of the necessity for those trained in musical conservatoires to be experienced in performing under pressurised circumstances. My participants will have had to perform to peers, teachers, and musicians of high esteem on a regular basis. The idea that they will be judged based on their musical abilities is something that all trained performers face daily, and with which they are intimately familiar. Indeed, it is plausible that in the presence of an

²¹ Speech scientists also have this non-neutrality when recording speech.

audience—as opposed to being alone in the recording studio or practice room—the participants would find themselves more motivated to convey an expressive performance.

Another disadvantage of being an ‘insider’ relates to the extra care in which I must deal with the collected data. As I was often present and active within the conservatoire during the time of my research, any leak of data could potentially result in the identification of a participant. To prevent this, I took extra care when discussing and demonstrating my work in public. As part of the ethical guidelines of my institution, I obtained informed consent from my participants and guaranteed their anonymity.

Although any experiment requires expertise to be conducted properly, my determinant role throughout the design, elicitation, analysis, and interpretation of this study allowed for purely ‘practitioner-led’ research; this, to my knowledge, is unprecedented among the published work in this field. Furthermore, there is a difference between conducting a study on one’s own language versus a language that one has only meta-knowledge of; the same can be said for musical performance. Regarding the specificities of experimental design, my chosen musical excerpt, although previously unknown to the participants in that they had never seen or heard the piece, is clearly intelligible within the stylistic norms of the Western classical music canon. During the recording process, I gave instructions using vocabulary and imagery common within piano performance study, including the direction, ‘Please perform this excerpt expressively, as if you are on stage’; these and other decisions detailed above describe, I trust, the design of an ecologically valid study. Furthermore, as a performing musician, my training and experience in critical and analytical listening allowed for ease and accuracy during the annotation process of the musical performances. Regarding my empirical setting, I argue that my decision to conduct experiments at our common current musical institution brought

advantages: time efficiency, access to high quality facilities and equipment, and a range of potential participants.

I will end this section with a response to some reservations raised by Patel and Daniele (2003). They question the use of performed music for measuring rhythmic values such as nPVI. Although they concede that linguistic nPVI values are measured from acoustic sources, using actual musical performances meant facing certain questions that were difficult to answer, including which performance of the piece is to be measured, and how is this performance deemed more suitable than other performances? Furthermore, each performance will differ in precise timing of notes, affecting the resulting data. My answer to these concerns is to cite the 'ambiguous' nature of musical performance that Patel and Daniele sought to avoid, and conduct a proper and valid comparison between performed music and speech. The concerns that Patel and Daniele have voiced regarding performance variation in music can also apply to the 'problem' of performance variation in spoken language; that is, timing in speech is dependent not only on phrase structuring and prosodic emphasis, but also on rate of speaking and other speaker-related dimensions. There may be quite different rhythm metric scores for the same language depending on speaker and speaker-style (as Arvaniti, 2009, among others, have stressed), thus the question of which performance of the spoken language is to be measured also applies. My very particular position has allowed me to control for both the speakers and performers of my speech and musical data, giving my study an advantage over others. Finally, my perspective as a classical musician provides me with extensive experience in the analysis of subtle differences in timing between performances (either of my own recordings or of those of others), as it is one of the ways in which musical performers learn and develop. I came to realise at the start of my research journey that analysing my own performances provides only a partial

view; that in order to answer my specific research questions, I had to broaden the scope of observation to include other performers and their music.

3.3 Participants

A total of eleven classical pianists participated in my study, six of whom were also involved in the pilot. Five pianists were native speakers of Mandarin Chinese with varying levels of L2 English, and the remaining six were monolingual speakers of Standard British or General American English.²² At the time of the recordings, ten participants were students at my conservatoire and one participant had graduated the year prior but was, at that time, employed there as a staff pianist. Five participants were undergraduates, and six were postgraduates or had completed their postgraduate study. My requirements for the participant selection were as follows:

- i) pianists had to be trained in the Western classical tradition and either studying for or had obtained a degree in piano performance at a conservatoire;
- ii) the native languages of the pianists had to be either Mandarin Chinese or English;
- iii) the native Mandarin speakers had to have English as second language (L2) exclusively; and
- iv) the English speakers had to be monolingual.²³

²² While speech rhythm may differ between dialects within each language group, I decided that the two chosen languages for this study were contrasting enough for rhythmic differences between languages to be clear enough regardless of dialectal variation.

²³ One of the English monolingual participants (Eng 5) had some exposure to Mandarin as a small child but has been functionally monolingual in English since age 5. This participant identifies predominantly as British, with a little knowledge of Chinese culture.

Gender was not a variable under investigation in my research, or controlled for, but for reference, four of the native Mandarin speakers were female and one was male, while five of the English monolinguals were male and one was female.²⁴ In the questionnaire, the English monolinguals self-identified as belonging to one of two categories, 'British' or 'American'. The Mandarin speakers identified themselves as strongly 'Chinese', equally 'Chinese and British', or 'Chinese and American'; this seemed to depend partly on how long they had lived abroad and where, and partly on their own sense of self-identity.

The participants were recruited through personal association; all were peers who expressed interest in my study during informal conversations, either in person or through personal online messaging. There was no payment for participation, and the scheduling of recording sessions was made according to the availability of the participants. All participants were given an information sheet, detailing the process and general aims of the study, and a consent form (see Appendix A). A questionnaire to establish language ability and background was also given (see Appendix C). All participants were made aware that they could stop at any point during the study, should they feel uncomfortable or become unwilling to continue. The participants were also guaranteed anonymity and were assigned a number, along with their language group, for reference throughout the study (e.g. Mandarin 1, English 2).

²⁴ I acknowledge that the gender imbalance is at least potentially significant; further study should seek to control for gender as an influencing factor.

3.4 Criteria for Selecting Materials

Music

My criteria for choosing the appropriate musical excerpt to be performed involved the following four characteristics:

- i) Simplicity
- ii) Expressive lyricism
- iii) Unconventional tonal language
- iv) Unfamiliarity

The first requirement was necessary for successful learning and performance; simplicity of composition allowed for the participants to learn and perform the excerpts without difficulty and error. Additionally, the musical material had to be simple in metre and texture to allow for ease of annotation and analysis (i.e. involving unambiguous note and beat onsets and offsets).

The second requirement relates to music that can be described as ‘lyrical’, ‘melodic’, and ‘expressive’ in character. Lyrical melodic material calls for ‘vocal-like’ phrasing and expressive nuance in its performance, exactly the elements that my study was targeting. This quality of ‘singing’ phrases has its roots in 18th-century instrumental performance treatises. Sources typically advise instrumentalists (particularly keyboardists) to learn and understand the essence of expression in performance by listening to artistic singing (e.g. Türk, 1789, in Doğantan-Dack, 2012). Today, this concept continues to have a strong influence on expressive musical performance.

The third requirement sought to control for the presence of conventional harmonic organisation within the music that may suggest to the performer certain expressive

decisions. It was important to choose a piece featuring a use of harmony that departed from so-called 'common practice' norms and therefore limit conventional interpretive responses, so that participants would rely more on instinctive and individual expression during performance. These conventional responses (e.g. requiring more emphasis and thus longer duration of certain chords due to their (musically) syntactic relationships, such as in a V-I cadence) would necessarily result in less variation between performers (all of whom had undergone similar training in the Western classical music tradition).²⁵

The final requirement related to the participants' familiarity with the composition; it had to be a piece that the participants had no previous knowledge of. This was intended to guarantee that the participants' performances were not the object of any external influences, whether this was through knowledge of performances by others or having received instruction on the piece by a teacher. By including this requirement, I controlled for the personal ownership of the participants' expressive decisions; I could also ensure that the recorded performances were representative of an early stage in a performer's learning of a piece, where the expressive decisions are, in a sense, raw and unfiltered.

For my experiment, I chose the first two lines of 'Rosemary', a short piece for solo piano from Frank Bridge's *3 Sketches*, H. 68 (1906). Bridge was an English composer, violist, and conductor, and was active during the first half of the 20th century. 'Rosemary' is one of his early works, written during the first decade of the century. The majority of his compositions during this period are vocal and chamber music. Notably, works such as the

²⁵ Of course, it is still be possible for a piece with conventional harmonic organisation to result in differing interpretations, but this seems to correlate with the differences in expertise and experience of the performers (Palmer, 1989).

String Quartet, No. 1 (1906) and the Phantasy Piano Quartet (1910) are said to have ‘a fresh, distinctive lyric impulse’ (Griffiths and Dibble, 2001). Sources mention the influence of Charles Villiers Stanford, Bridge’s composition teacher from 1899 to 1903, upon the conservative nature and lyricism of his early work (Payne et al., 2001).²⁶

My decision to choose an English-language composer’s music relates to my focus on English speech in my research. As studies have suggested that both the speech rhythm of English and the musical rhythm of English composers’ themes may have high vowel/note durational variability, I believe it is fitting to use English music to compare durational variability between performances.



Figure 2: 'Rosemary' by Frank Bridge (mm 1–8)

²⁶ Stanford was best known for his vocal and liturgical compositions; his melodies were often infused with the ‘contours of Irish folk music’ (Dibble, 2001).

Speech

The speech material consists of two excerpts, one from an English story (*A Serious Case*; Rose, n.d.) and one from a Mandarin Chinese story (窗外 or *Outside the Window*; Learn Mandarin Chinese, 2015). These excerpts were presented in typed format to the participants. Both stories were taken from language-learning websites and are of beginners' level in difficulty. Figure 3 shows the English excerpt; for the Chinese story, see Appendix D.

I have a friend who is afraid of spiders.

This isn't very unusual; a lot of people are afraid of spiders. I don't really like spiders much myself. I don't mind them if you see them outside, in the garden, as long as they're not too big. But if one comes in the house, especially if it's one of those really big spiders with furry legs and little red eyes, then I go "yuck" and I try to get rid of it.

Figure 3: 'A Serious Case' (Rose, n.d.)

3.5 Pilot Study

The following details my pilot experiment, including the materials, instrumentation, experimental procedure, and findings. It was necessary to conduct a preliminary study to determine an efficient and effective procedure, as well as to establish the dependent and independent variables of my main experiment. The pilot also revealed certain practical issues, relating to the amount of materials used for testing, as well as the targeted variables.

Materials

My pilot involved recording the performances of two musical excerpts, one of Frank Bridge's 'Rosemary' and the other of 'Yu Diao' (1984), from *Two Chinese Bagatelles*, written by the Chinese composer Chen Yi (1953-). My aim of having two excerpts was to represent English and Mandarin in both the speech and musical material.



Figure 4: 'Yu Diao' by Chen Yi, mm 14–18

Dr. Chen Yi was born in China and studied music composition at the Central Conservatory in Beijing. She received her Doctor of Musical Arts degree in composition from Columbia University. Her work is marked by different musics and aesthetics, including Western classical, Marxist-Maoist, Chinese traditional, and Western contemporary music (Li, 2003); however, she considers her music as belonging to the Western classical tradition (Lai, 1999). 'Yu Diao' is a solo piano piece, written for children. The title, as explained by Chen Yi, translates to 'tune of the Henan Province' (Li, 2003). Chen Yi was influenced primarily by the local operas of Henan province, and used pitch material from these sources to create the melodies in 'Yu Diao'. The excerpt chosen for my experiment is the second theme (mm. 14-18), which is described as 'gentle, smooth, and lyrical' (Li, 2003).

The two musical excerpts were recorded in two different versions: with and without expression. The second, 'mechanical' version required that the performer make no interpretative nuances to her or his performance. This version served as the baseline for each participant's performance of the excerpt; the subtle changes in timing and *rubato* that occurred in the expressive version could be compared to the mechanical version, thereby isolating and identifying the changes as expressive decisions that are particular to each individual.

The speech material consisted of a) five English sentences (sourced from Ramus, Nespor, and Mehler, 1999), b) an English story ('A Serious Case', see Figure 3), and c) for Mandarin speakers only, two Chinese stories – one at a beginner's reading level (*Outside the Window*) and the other, advanced (*Traces of Sin 1*; Learn Mandarin Chinese, 2015). These were presented to the participants in typed form. Additionally, the participants held a recorded conversation with the researcher; they were asked to describe what they had for breakfast that day, how they travelled to the conservatoire, and what the weather was like. The conversations were recorded as spontaneous speech samples, in contrast to the read speech.

Experimental Procedure

Music

- i) The participants were presented with the two musical excerpts and speech materials along with the information sheet and consent form. They were asked to read the information and, once any issues were clarified, sign the form. This was done four to six days prior to each recording session, potentially giving the

participants a modest but sufficient amount of time to prepare for performance.

This, as opposed to presenting all participants with the music at the time of recording, sought to prevent errors in performance and any discomfort due to sight-reading. Note: the exact amount of practice time for each participant, prior to the recording session, is unknown. The quality of performance during the recording was based on the participant's own judgment. This reflects the preparation process of a performing musician in real-world settings, and is highly individual.

- ii) Participants took part in the recording sessions individually, in the presence of the researcher (myself) and a member of the conservatoire's recording staff. All staff agreed to confidentiality to protect the identities of the participants.
- iii) Participants were asked to prepare two versions of each excerpt, one without expression, i.e. 'mechanical', and one with expression, 'as if on stage, in concert'.
- iv) During the recording session, all participants were given the opportunity to practice on the piano used for recording, at their discretion. Note: practice was varied between participants; some spent up to, but no more than, four minutes, while others played the excerpt through only once. Some did not wish to practice at all.
- v) The participants were directed to play the 'mechanical' version of the 'Rosemary' excerpt by following the notated durations and articulation on the page without adding any interpretation of their own. They could use the sustain pedal, but only for note-length consistency (certain large intervals were impossible for participants to play *legato* without the aid of the sustain pedal).

- vi) After a brief pause (during which they were encouraged to practice, if they wished, in the new mode), the participants played the excerpt in an expressive manner. Note: If there were any accidental wrong notes/rhythms or errors in note/rhythm reading, the participants could re-record as many times as they deemed necessary. Judgment of expressiveness was made exclusively by the participants themselves. Accuracy judgments (notes and rhythms) were made jointly by the participants and myself.
- vii) Steps v) and vi) were repeated for the 'Yu Diao' excerpt.

Speech

- i) Following the musical performances, the participants entered a vocal booth to record the speech materials.
- ii) The participants were presented with the typed page of English text and given some time to read through before recording.
- iii) Once prepared, the participants read through each portion of the text (first the English sentences, then the story) using their own chosen pace and manner. Recordings were stopped in between each style of text.
- iv) Once the texts were recorded, the participants were asked to drop the piece of paper and engage in a conversation with the researcher.
- v) For the Mandarin participants, steps iii) and iv) were repeated with the two Mandarin stories; the spontaneous-speech conversation was made on the same topics as in English.

Instrumentation

The six pilot recording sessions were made in the recording studio of the conservatoire. The piano was a Steinway Model B grand piano and two DPA 4011 (cardioid) microphones were used to record the performances. The speech recordings were made in a vocal booth, equipped with a Neumann U87 microphone. Both music and speech were recorded using ProTools® software.

Annotation

The recordings were trimmed with the 2.1.1. version of the Audacity® software. Speech recordings were visualised and annotated through PRAAT (Boersma and Weenink, 2018). Each sentence was annotated in three tiers: word, syllable, and C/V. Within the syllable tier, start-points and end-points of phrases (both final and intermediate) were labelled with (*); this was to facilitate the extraction of durations according to position in the phrase. The vowel and consonantal intervals were segmented based on auditory, waveform and spectrographic inspection, and segmentation was carried out with reference to the following standard criteria: placement of boundaries between vocalic and consonantal intervals were guided by a sudden, significant drop in amplitude and a break in formant structure, particularly F2. The marking of consonant onsets was facilitated by various cues, according to the manner of the consonant. In addition to syllable boundaries, syllables were also labelled according to segmental content and according to level of prominence. For prosodic prominence, syllables were labelled as belonging to one of three levels: i) unstressed (u); ii) stressed (s); iii) nuclear stressed (SS). Both intonational and intermediate phrase boundaries were identified and marked. Furthermore, syllables, as well as vocalic

and consonantal intervals, were categorised according to phrase position (initial, medial, or final).

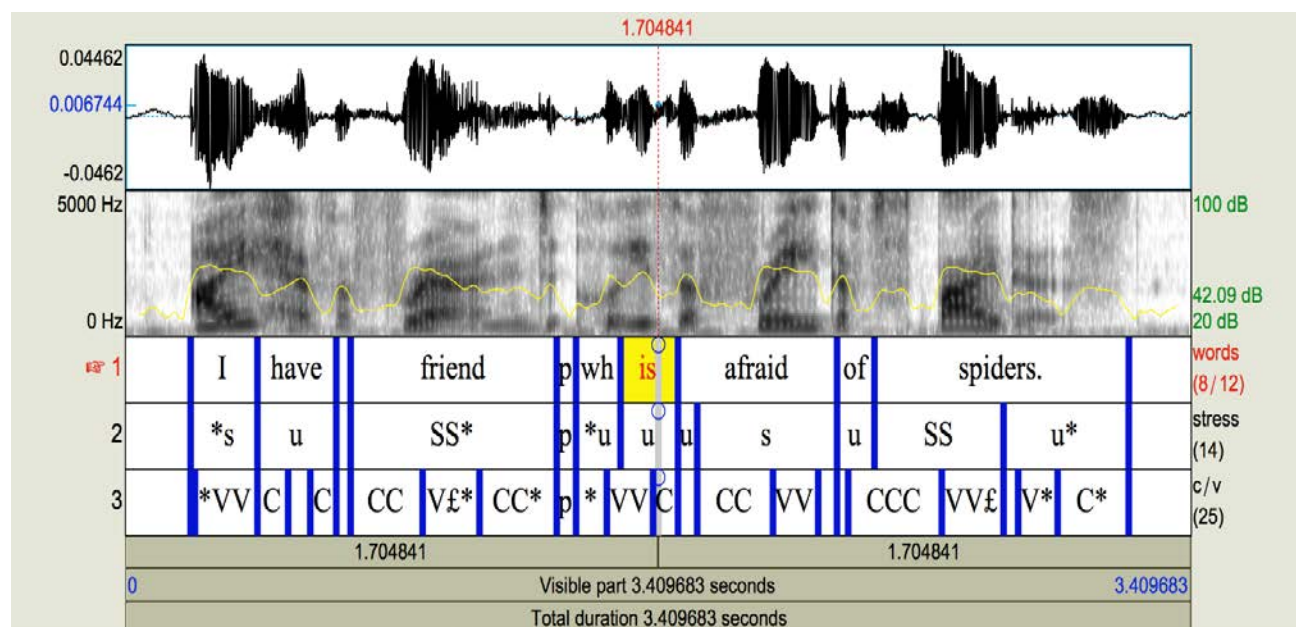


Figure 5: Textgrid, waveform, and spectrogram of Mandarin 1, sentence 1

Spectrographs of the performances of ‘Rosemary’ were generated by Sonic Visualiser; this application allows for the visualisation, annotation, and data extraction of audio music recordings. Similar to the speech material, labelling of the performances was aided by spectrographic and auditory information. For the expressive performances, timings of beat onsets were labelled on three separate beat levels (‘time instance layer’): i) semi-quaver beat; ii) quaver beat; iii) crotchet beat. Bar timings were also labelled, in a separate layer, for ease of reference between the musical score and the annotations. A curved line graph, made of the time instances in the quaver beat layer, was added in a separate layer (‘time values layer’). This shows, visually, the lengthening and shortening of quaver beats (*rubato*) throughout the excerpt; the peaks of the curve show lengthening, while the troughs show shortening. Of the mechanical performances, annotations of quaver beat durations, as well as a curved line graph, were also made (see Appendix F). Furthermore, the durations of

each note of the melody (see the horizontal green lines in Figure 5) were annotated by setting a point at the onset of each note (the offset of the final note was also labelled). A timing curve was also made of the melody in each performance.

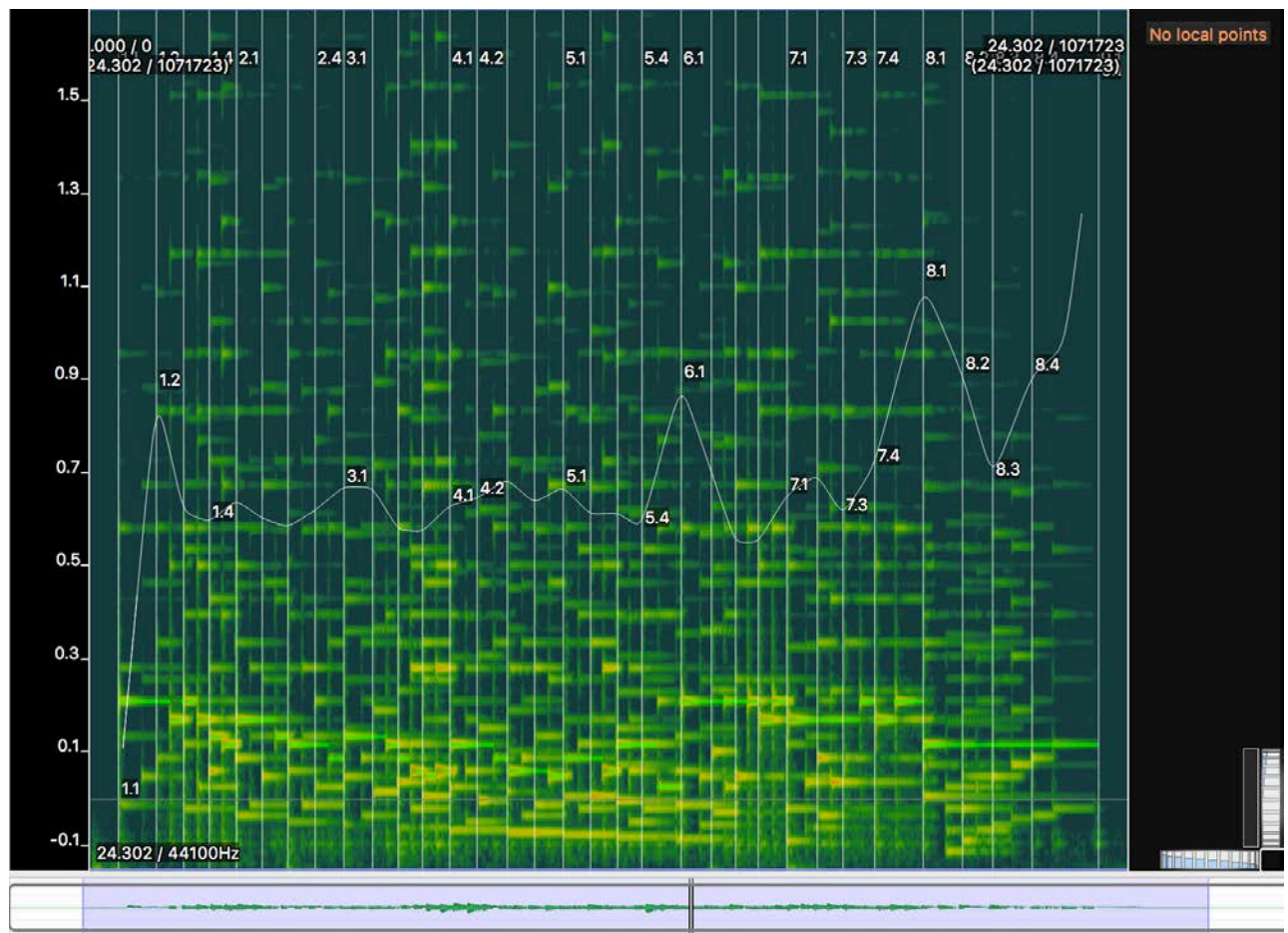


Figure 6: Spectrogram of expressive performance by Mandarin 1, with quaver beats (vertical white lines) and quaver beat timing curve (horizontal white curve)

Discussion and Variables

The pilot experiment allowed me to refine several elements of my methodology, relating to materials, procedure, and variables. First, I decided to annotate only the recordings of the English piece and omit the recording of the Chinese composition altogether from the final experiment. This decision stemmed from a comment from one of the Mandarin-speaking participants during a recording session: as a response to my direction to ‘perform this music

as if you're on stage, giving a concert', the pianist lightly mentioned that she had never performed 'Chinese music' in a concert before. I was thus reminded of the bias within Western classical music training with regards to taught repertoire: higher education music institutions tend to place focus on 'orthodox' repertoire, which mainly involve compositions from Western (usually Caucasian and male) composers and dated not later than the mid-20th century. The works of even well-established composers from East Asia, most of whom were trained either in the West or within the Western tradition, have not found a place in the orthodoxy of classical music. As all my participants were trained in the Western classical tradition, the performing of Chinese compositions would not be a familiar and natural task. I felt this posed a risk to the ecological validity of the experiment, so I decided to omit the Chinese excerpt in future experiments. The second refinement involves the exclusion of several speech materials, namely the simple English sentences, the advanced Mandarin story, and both English and Mandarin spontaneous speech, from my annotations and subsequent experiment. They were excluded due to their lack of comparability to the musical material. The simple sentences were random in content and without context, while the musical excerpt (refer to Figure 1), taken directly from the published score, had a clear stylistic context. The spontaneous speech material would be better compared to improvised musical performance, rather than performance of notated music. Ultimately, I decided that the narrative speech material (the English story and simple Chinese story) was most comparable to the musical material.

I also realised that I had to be more flexible with the scheduling of time and place of elicitation due to the busy schedules of my participants. Although I had originally sought to be consistent with the instrumentation (microphones, piano, and recording system) used for all participants, this was not a practical objective as almost all the participants were full-

time students with demanding schedules of classes, performances, and/or personal practice. Consequently, I had to change the location and instrumentation for all post-pilot recording sessions to allow for ease of scheduling; I booked practice rooms with grand pianos of similar size and model as the piano used in the pilot (Yamaha S6 and Steinway B models), and utilised a portable recording device that is popular among performers due to its practicality and transportability (Zoom® H4n Handy Recorder).

Finally, I eliminated certain types of measurements for extraction and analysis. Of my original speech annotations, I decided not to measure for phrase-final lengthening or prosodic prominence. Of my musical annotations, only the durations within the quaver beat level were extracted—crotchet and semi-quaver beat durations were omitted, along with the melodic note-durations. The first elimination came down to matters of time and scope; including all the other measures in this research would have been simply unmanageable. As for the omission of certain beat-levels, I discovered through annotation that the timing curves at the quaver beat level demonstrated sufficient differentiation between the participants (see Appendix F for timing curves of all performances); crotchet-level timing curves showed less detail (due to larger spans of timing points, or ‘instances’) and were therefore too similar among participants for proper comparison; semi-quaver timing curves were too detailed (with too many time instances), making comparisons tedious. Omission of melodic note-durations relates to my interest in the variability of *rubato*, rather than note durations between performers, as discussed in Section 3.2.

3.6 Main Experiment

For my main experiment, I recruited a further five participants (two native Mandarin speakers and three English monolinguals) to record the excerpt of ‘Rosemary’ as well as the

English and Mandarin stories. The recordings took place in three separate practice rooms, on grand pianos (one room had a Yamaha S6 and the other two had Steinway B models). The Zoom Recorder was placed near the piano during performances and near the participant during the speech recordings. The English and Mandarin stories were presented to the participants on a typed sheet. The procedure of this experiment followed that of the pilot, except for the omissions of the Chinese composition, the English speech sentences, the advanced Mandarin story, and the spontaneous speech.

Here follows the list of variables established during the pilot but refined in the main experiment; in addition to certain rhythm metrics, I included speech rate and tempo as a dependent variable in their respective categories, as I was curious to see if there are any correlations between the rate of one's speech and one's performance tempo.

Variables

Speech

Independent variable: native language (two levels: Mandarin Chinese; English)

Dependent variables:

- i) %V = the proportion of vocalic sounds in a given utterance
- ii) ΔV (standard deviation of vocalic interval duration), nPVI_V (normalised variation of duration in successive intervals), VarcoV (normalised standard deviation of vocalic duration): the durational variability of vocalic sounds
- iii) ΔC (SD of consonantal interval duration), VarcoC (normalised SD of consonantal interval duration): the durational variability of consonantal sounds
- iv) speech rate: syllables per second

Music

Independent variable: performance style (two levels: expressive; mechanical)

Dependent variables:

- i) Δ , nPVI, and Varco: the variability of quaver beat durations
- ii) overall tempo: quaver notes per minute

Possibly Confounding Variables Not Controlled For

- i) time spent in native country
- ii) time spent in an English-speaking country
- iii) age at which English learning started, and time spent learning English (for Mandarin speakers)
- iv) (self-identified) culture
- v) gender
- vi) musical performance experience

3.7 Ethics and Limitations

As my study involves comparative analysis on the performances (both in speech and music) of people from different cultural and racial backgrounds, there was a possibility of risks of an ethical nature. These risks relate to the participants' own comfort, security, and reputation as well as to more general issues that may arise from my findings. A significant issue was the possible misconstrual of my study as being in support of racial and cultural prejudices. As mentioned briefly in Chapter One (and to be further discussed in Chapter

Five), my research is certainly not intended to validate or further entrench East/West stereotypes in classical music. However, it is a plausible concern that certain findings could be misinterpreted as suggesting that one group of musicians performs ‘more expressively’, or in some way ‘better’. To avoid this, I had to be extremely careful, initially with my methodological approach, and subsequently with my experimental design and process.

There are two significant ethical concerns, particularly within ethnographic research, that are further magnified by the presence of cross-cultural and/or ‘insider’ factors: representation of truth and confidentiality. As Ford (2010: 94) has pointed out (in reference to representation within research on music in higher education), the music world is small and made up of ‘close social networks where everyone seems to know each other’ (see deNora, 2000). This, combined with the idea that ‘people tend to stake their identity on their musical practices’ (Ford, 2010; see Schneider, 2010), means that great pressure is placed upon representation to accurately depict ‘the truth’ (which, in itself, is a source of debate, see Ford, 2010: 93–94). As an insider, I was particularly aware of the potential dangers of misrepresentation, initially striving for a qualitative, ethnographic methodology that could take proper account of the culture, musical background, and personal experiences of the participants. However, this meant a plethora of ethical questions, including accuracy of representation (‘Am I presenting this participant in the most accurate and truthful light, void of my own judgments and interpretations?’). Ultimately, my experimental research sought to avoid possible generalisations with regard to culture and cultural dichotomy by focusing on quantifying and comparing two elements (speech and music) existent in all cultures that share a common framework. Furthermore, I defend my decision to use a quantitative method for investigating cross-cultural musical expression,

which is typically associated with qualitative means of inquiry: I maintain that there is great benefit in providing different ways of observing the empirical world.

My methodology also sought to alleviate the issue of potential reputational damage for the participants, both within the conservatoire and in public. I refrained from including detailed profiles of each participant in this thesis, as information about familial, educational, and self-described cultural background could reveal the identities of the participants. Fortunately, within any given musical conservatoire in the U.S. and U.K., there are large-enough numbers of piano students who are native Mandarin speakers, as well as English monolinguals, to make identification based on language background improbable.

As mentioned earlier, in the section on musical materials, one of the stipulations for choice of music was unfamiliarity: participants must not have known of the piece prior to participation. Given the short amount of preparation time, as well as restrictions against external influences on expressive decisions, the participants were effectively presenting an initial interpretation of the music that is particularly individualised and instinctive. Under ‘normal’ circumstances, the music that a performer presents on stage would have gone through several cycles of transformation due to external factors such as exposure to other performances of the music, advice (musical and/or practical) from others, or simply with the passing of time. Having this knowledge as a practitioner, I was assured enough to conduct an experiment on pianists’ performances without being overly concerned for how the results would be judged as a representation of their overall musical and expressive abilities. Again, my study was more interested in the resulting raw and unfiltered expressivity of the pianists, rather than in their most accomplished and presentable work.

Naturally, this study contains limitations, as certain variables could not be fully controlled. For instance, because I collected my data from a music conservatoire in London,

I was not able to recruit completely monolingual speakers of Mandarin to compare with the English monolinguals, although this in many respects would have been ideal. However, I justified that the chances of finding conservatoire-level classical pianists who could only speak Mandarin and who have never been exposed to English are slim, particularly in a world where information is easily-accessible through the internet and English-language media and pop culture are globally prevalent and influential. To compensate, I selected native Mandarin participants that encompass a wide range of exposure to the English language; for example, some participants had been living in the U.K. for only two years, and had little confidence in speaking English, whereas another participant had been living in English-speaking countries (the U.S. and England) for a total of six years (see Appendix C for participant profiles). Additionally, among the English monolingual participants, there were two national dialects, American and British English (excluding that of Wales and Scotland). Although it has been suggested that different dialects of English have differences of rhythm metrics in both speech and music (see McGowen and Levitt, 2010 for a study on the spontaneous speech and fiddle music of musicians from three distinct regions: Kentucky in the United States, Donegal in Ireland, and the Shetland Islands in Scotland), they are generally similar enough to be compared, together, with a language as dissimilar as Mandarin Chinese.²⁷

Beyond language, I had difficulty controlling for the amount of performance experience of my participants. Although it could be assumed that those who were younger and were undergraduates had less performance experience than the postgraduates, it is not

²⁷ Note that the study of McGowen and Levitt (2010) involves spontaneous rather than read speech, as well as music that has been acquired by ear, without a score.

always the case; some younger students could have performed a significant amount and competed in many competitions prior to entering a higher education institution.

Unfortunately, I did not gather enough information on the participants' musical background to form profiles of their performance experience. Neither was I able completely to control for—or even determine—the exact amount of time spent learning the excerpt. I could have chosen to set strict requirements for the amount of preparation time; however this would have raised other confounds, since individual performers learn scores and absorb musical information at varying speeds due to differing abilities in sight-reading, learning experience, and other factors such as life distractions, other musical responsibilities, and the amount of work or stress in their lives. Thus, the amount of preparation for this study was mainly self-governed, according to what felt right for each participant.

3.8 Analysis and Results

In this section, I present the results of my main study, combined with those relevant from the pilot. The following rhythmic measures were calculated for four English sentences (averaged for each person), as read by each speaker, and four Mandarin sentences (averaged for each person), as read by each Mandarin speaker: %V, ΔV , ΔC , VarcoV, VarcoC, and nPVI_V. Mean values and speech rates were also calculated for each speaker. For the durations of each musical performance (both mechanical and expressive), the mean, standard deviation, nPVI, and Varco measures, as well as overall tempo, were calculated. Due to time constraints, durational values for phrase-final units in both speech and music analysis were excluded from the calculations and reserved for future examination. In

speech, both pre-final and final phrase syllables were excluded; in music, penultimate and ultimate bars were excluded for the same purpose.²⁸

Speech rate (syllables per second or sps) was calculated manually by dividing the total number of syllables by the total length of the sentence. Again, the final and pre-final syllables were excluded to control for phrase-final lengthening. Large pauses (often triggered in speech by the presence of commas and semi-colons, and demarcating phrase boundaries) were excluded from the total length.²⁹ The scores of four sentences were averaged for each person.

In music, the average quaver-note duration was calculated by dividing the total duration by the total number of quaver beats; in this study, only 16 out of the 32 total quaver beats in the excerpt were analysed. This is because in calculations of performance tempo,³⁰ it is necessary to exclude beats that were involved in phrase-final lengthening and prominence cues (conspicuous slowing down, lengthening and/or delaying of notes to

²⁸ While final lengthening is a known difference between languages, and is assumed to be a factor of the rhythm percept, I was unfortunately unable to examine the data for this study and hope to do so in future work.

²⁹ Interestingly, three of the Mandarin speakers of English had a number of small pauses between non-boundary syllables, and it was decided that these should be included in the total length sum, as such pauses are interpreted as a characteristic of the three Mandarin speakers (when speaking L2 English).

³⁰ In retrospect, it might be more revealing and perhaps desirable to branch away from the common musical practice of tempo calculation for future investigations.

highlight importance).³¹ From the average quaver-note duration, the average metronome speed was determined (quavers per minute or qtm).

Results and Statistical Analysis

The means of %V, nPVI_V, VarcoV, VarcoC, and speech rate for four English and four Mandarin sentences (for those participants with native Mandarin) were calculated for each participant. A Mann-Whitney *U*-test was applied to the scores between language groups to test for significant difference between groups. The measures of nPVI and Varco for quaver beat durations of the musical performances were also tested for significant difference between language groups. In addition to testing for statistical significance, I calculated the Pearson product-moment correlation coefficient (PMCC) to determine the statistical correlation between the speech and musical performances of individual participants.

The following subsections present and discuss the results of measures. Vocalic variability measures in speech (nPVI_V and VarcoV) are presented together for explicit comparison between language groups. Results of VarcoC and %V in speech are paired as it is suggested that such measures are better able to demonstrate difference between Mandarin-accented English and native English, as well as between language groups (Mok and Dellwo, 2008). Speech rate and overall performance tempo are paired for explicit comparison between speech and music.

³¹ The 16 quaver beats were determined on an individual basis; each participant displayed prominence cues at slightly different moments. Judgements were made by myself.

nPVI_V and VarcoV in Speech (English and Mandarin)

Studies have found differences in measures of temporal variability—considered to be related to the percept of rhythmic difference—between English and Mandarin Chinese speech, with English having a higher measure of pairwise variability of vocalic interval durations (Grabe and Low, 2002; Benton et al., 2007; see also Li and Post, 2014), as well as for standard deviation of non-successive vocalic interval durations (Ramus et al., 1999), than Mandarin speech. However, the findings of these studies do not give conclusive evidence for the separation of Mandarin and English into distinct rhythmic groups; significant speaker-based and style-specific variation within groups support the theory that differences in the rhythmic percept are continuous, rather than categorical (Dauer, 1983; Dellwo et al., 2004; Dellwo, 2006). Nonetheless, the two languages are reported as sounding rhythmically distinct to listeners, and due to the typical transfer of L1 prosodic features onto L2 speech, this is plausibly also true for the L2 English of L1-Mandarin speakers, when compared with L1 English speech. However, application of rhythmic measures to L2 English (with L1 Mandarin) and L1 English speech have been inconclusive, with the nPVI_V measure being inconsistent of hypothesised language-based distinctions (Mok and Dellwo, 2008; Li and Post, 2014); the results of my study further address this inconsistency.

Table 1 shows the results of my analysis of the nPVI_V and VarcoV measures in English speech, for both L1 and L2 speakers of English:

Table 1: mean nPVI_V and VarcoV values for L1 and L2 English speech

Participant	nPVI_V	VarcoV
Man 1	65.96	50.38
Man 2	44.22	45.49
Man 3	32.78	33.88

Man 4	40	36.67
Man 5	49.03	46.75
Eng 1	46.71	40.08
Eng 2	52.22	49.43
Eng 3	50.84	45.43
Eng 4	50.14	50.69
Eng 5	58.19	63.29
Eng 6	50.9	55.77

The Mann-Whitney U -test determined that there was no significant difference between language groups for both nPVI_V and VarcoV measures of L1 and L2 English speech (both at $U = 7$, result *not* significant at $p < 0.05$). This seems to reflect the results from Mok and Dellwo (2008), which determined that Mandarin-accented English speech has a high nPVI_V value, similar to that of native English speech. Further examination of the results, however, show that Man 1 and Eng 1 have abnormally high and low scores, respectively; if these are excluded, a possible difference between L1 and L2 English may occur.³²

Table 2 shows the results of the native Mandarin speech, along with L1 and L2 English:

Table 2: mean nPVI_V and VarcoV values for L1/L2 English (columns 2 and 3) and L1 Mandarin speech (columns 4 and 5)

Participant	Eng nPVI_V	Eng VarcoV	Man nPVI_V	Man VarcoV
Man 1	65.96	50.38	53.47	45.34
Man 2	44.22	45.49	52.02	42.44
Man 3	32.78	33.88	41.34	34.78

³² This is simply predicted and not tested because, with the exclusion of Man 1, it was not possible to calculate for statistical difference as the minimum number of samples for each group cannot be under 5.

Man 4	40	36.67	37.4	34.98
Man 5	49.03	46.75	45.50	38.57
Eng 1	46.71	40.08	-	-
Eng 2	52.22	49.43	-	-
Eng 3	50.84	45.43	-	-
Eng 4	50.14	50.69	-	-
Eng 5	58.19	63.29	-	-
Eng 6	50.9	55.77	-	-

Comparisons between L1 English and L1 Mandarin speech revealed no significant difference between language groups for the nPVI_V measures ($U = 9$, result *not* significant at $p < 0.05$). However, there is a significant difference between the VarcoV measures of the L1 languages ($U = 2$, result significant at $p < 0.05$), with higher VarcoV scores for English. Interestingly, even when Eng 1 (who shows lower variability than her/his L1 English peers) is excluded from the calculations, nPVI_V measures still fail to show a significant difference ($U = 7$), while VarcoV scores show highly significant difference between language groups ($U = 0$). This suggests a lesser degree of variability in L1 Mandarin, at the global level, than in L1 English, while no apparent difference at the localised level.

Interestingly, comparisons between L2 English and L1 Mandarin revealed no significant difference for the nPVI_V measures ($U = 11$, result *not* significant at $p < 0.05$) and VarcoV measures ($U = 8$, result *not* significant at $p < 0.05$). However, upon further observation, Man 2 and Man 3 seem to be outliers for nPVI_V; the other speakers have higher scores in L2 English than in L1 Mandarin, while Man 2 and Man 3 have lower scores. For VarcoV, only Man 3 has a lower score in L2 English than in L1 Mandarin.

VarcoC and %V

In Mok and Dellwo (2008), measures of VarcoC and %V were better able to show a difference between Mandarin-accented English and native English speech than the vocalic variability measures: VarcoC was shown to be higher in L1 English speech than in Mandarin-background L2 English; conversely, %V was lower in L1 English speech. This is interesting because consonant measures and %V are known to be more reflective of phonotactics (the linguistically specific composition of vowels and consonants in syllables); thus, differences in such measures suggest that the Mandarin speakers of English in Mok and Dellwo's study may have been changing the English syllables in their speech, perhaps due to insufficient fluency, e.g. by simplifying clusters, or deleting codas. In Li and Post (2014), consonantal metrics between native Mandarin-, German-, and English-speaking language groups (measuring L1 Mandarin, German, and English speech, as well as L2 English speech) were calculated from CVN (consonant-vowel-nasal), mixed, and CV-type sentences, to control for effects of syllable structure. Results suggest that consonantal metrics are indeed sensitive to syllable structure, as the effects of consonantal variability between language groups disappeared in CV-type sentences. Thus, differences in metric scores could be interpreted as relating to pronunciation difficulties of consonants within more complex syllables (Li and Post, 2014: 246).

Table 3 shows the results of these two measures in my data:

Table 3: mean VarcoC and %V of L1 and L2 English, and L1 Mandarin speech

Participant	Eng VarcoC	Eng %V	Man VarcoC	Man %V
Man 1	60.33	43.9	44.85	47.69
Man 2	59.97	44.3	48.07	51.1
Man 3	55.79	51.4	44.97	57.05

Man 4	64.36	43.95	66.93	49.5
Man 5	55.69	48.25	59.04	55.54
Eng 1	64.52	48.35	-	-
Eng 2	62.39	43.33	-	-
Eng 3	61.46	46.33	-	-
Eng 4	63.7	48.43	-	-
Eng 5	63.24	43.63	-	-
Eng 6	60.3	41.45	-	-

The Mann-Whitney U -test determined a lack of significant difference between the two language background groups for both VarcoC ($U = 6$, result *not* significant at $p < 0.05$) and %V ($U = 11$, result *not* significant at $p < 0.05$) of L1 and L2 English speech. However, there does seem to be moderate intraspeaker correlation between VarcoC and %V; calculation of the PMCC shows a moderate negative linear relationship via a fuzzy-firm linear rule ($r = -0.4$). This suggests that speakers with higher VarcoC have a lower vocalic percentage; inversely, speakers with lower VarcoC have a higher vocalic percentage.

Interestingly, although there is a lack of significant difference between the native English and native Mandarin measures of VarcoC ($U = 6$), there is a significant difference between the %V measures ($U = 2$); native Mandarin is shown to have a higher %V than native English speech. This is to be expected, given the greater preponderance of CV syllables in Mandarin, and comparative lack of complex clusters. Furthermore, comparisons between L2 English and native Mandarin speech showed no significant difference of VarcoC ($U = 7$) and %V ($U = 4$). These results may be disproportionately due to two participants: Man 4 and Man 5. Both Man 4 and Man 5 show an inverse of the expected relationship between their VarcoC and %V measures of native Mandarin speech; rather than having a lower VarcoC and higher %V (as expected for Mandarin speech), the two speakers have a

higher VarcoC and lower %V (as expected for English speech). More on this will be presented in later discussions.

Speech Rate and Overall Tempo

The following table (Table 4) shows the rate of English speech, in syllables-per-second, and the overall performance tempo, in quavers-per-minute.

Table 4: mean rate of English and Mandarin read speech (syllables per second) and overall tempo (quavers per minute)

Participant	English rate	Man. rate	Overall tempo
Man 1	4.17	5.05	96
Man 2	4.61	4.33	77
Man 3	4.87	5.49	100
Man 4	5.79	6.2	102
Man 5	5.35	6.09	81
Eng 1	6.32	-	117
Eng 2	6.5	-	85
Eng 3	5.91	-	80
Eng 4	5.74	-	96
Eng 5	5.88	-	93
Eng 6	5.34	-	107

The Mann-Whitney *U*-test determined a significant difference between the two language groups for speech rate, with L1-English speech having a higher speech rate than L2 Mandarin-background English ($U = 3$, result significant at $p < 0.05$). Furthermore, four out of the five Mandarin speakers showed a higher speech-rate in their native Mandarin than in their English speech. This is supportive of the common phenomenon of second-language

speakers to speak slower in their second language than in their first (Mok and Dellwo, 2008). There is no significant difference between the speech rate of native English and native Mandarin ($U = 9$, result not significant at $p < 0.05$).

For overall music tempo, there is no difference between the two language groups ($U = 12.5$, result not significant at $p < 0.05$). The PMCC shows a weak positive linear relationship via a shaky linear rule between the L1 and L2 English speech rates and overall tempi ($r = 0.16$). Correlation between the Mandarin speech rates and overall tempi is also weak ($r = 0.24$). Interestingly, correlation between the native English speech rates and overall tempi shows a weak, *negative* relationship ($r = -0.14$).

Correlations Between Rhythmic Metrics of Music and Speech

Table 5 shows the nPVI and Varco measures of the musical performances:

Table 5: mean nPVI and Varco scores of quaver beat durations

Participant	nPVI-m	Varco-m
Man 1	10.43	13.25
Man 2	11.76	12.57
Man 3	8.24	7.96
Man 4	10.14	10.89
Man 5	4.58	5.04
Eng 1	14.21	22.07
Eng 2	11.51	14.02
Eng 3	11.22	10.39
Eng 4	12.68	15.36
Eng 5	12.15	14.89
Eng 6	12.33	16.65

Application of the Mann-Whitney test to the Varco scores of the musical performances (Varco-m) revealed significant language-based grouping; Varco-m scores were higher in the L1 English group compared to the L1 Mandarin group ($U = 3$, significant at $p < 0.05$). Likewise, nPVI-m scores were higher in the L1 English group compared to the L1 Mandarin group ($U = 2$, significant at $p < 0.05$).

Table 6 provides a complete overview of the speech measures of %V, nPVI_V, VarcoV, and VarcoC, along with Varco and nPVI measures of the musical performances:

Table 6: mean %V, nPVI_V, VarcoV, and VarcoC of L1 English, L2 English, and L1 Mandarin speech; mean Varco and nPVI of music

Participant	Eng nPVI_V	Eng VarcoV	Eng %V	Eng VarcoC	Man nPVI_V	Man VarcoV	Man %V	Man VarcoC	nPVI-m	Varco-m
Man 1	65.96	50.38	43.9	60.33	53.47	45.34	47.69	44.85	10.43	13.25
Man 2	44.22	45.49	44.3	59.97	52.02	42.44	51.1	48.07	11.76	12.57
Man 3	32.78	33.88	51.4	55.79	41.34	34.78	57.05	44.97	8.24	7.96
Man 4	40	36.67	43.95	64.36	37.4	34.98	49.5	66.93	10.14	10.89
Man 5	49.03	46.75	48.25	55.69	45.50	38.57	55.54	59.04	4.58	5.04
Eng 1	46.71	40.08	48.35	64.52	-	-	-	-	14.21	22.07
Eng 2	52.22	49.43	43.33	62.39	-	-	-	-	11.51	14.02
Eng 3	50.84	45.43	46.33	61.46	-	-	-	-	11.22	10.39
Eng 4	50.14	50.69	48.43	63.7	-	-	-	-	12.68	15.36
Eng 5	58.19	63.29	43.63	63.24	-	-	-	-	12.15	14.89
Eng 6	50.9	55.77	41.45	60.3	-	-	-	-	12.33	16.65

These results suggest that, although language-based grouping is not conclusive in the speech measures between L2 English and L1 English speech, there *is* a difference in VarcoV and %V between L1 English and L1 Mandarin. As for the inconclusive results between L1

English and L1 Mandarin for nPVI_V, perhaps this relates to the exclusion of pre-final and final syllables in the measurements; further discussion below.

Following these observations, I tested for intraspeaker consistency between the speech and music measures. Results of the PMCC concluded that, between VarcoV of (L1 and L2) English and Varco-m, a weak linear correlation exists ($r = 0.28$). However, the graph (Figure 7) below shows two outliers: Man 5 and Eng 1. Results of the PMCC, excluding Man 5 and Eng 1, reveal a *strong* linear relationship between native and L2 English speech and musical performances ($r = 0.85$).

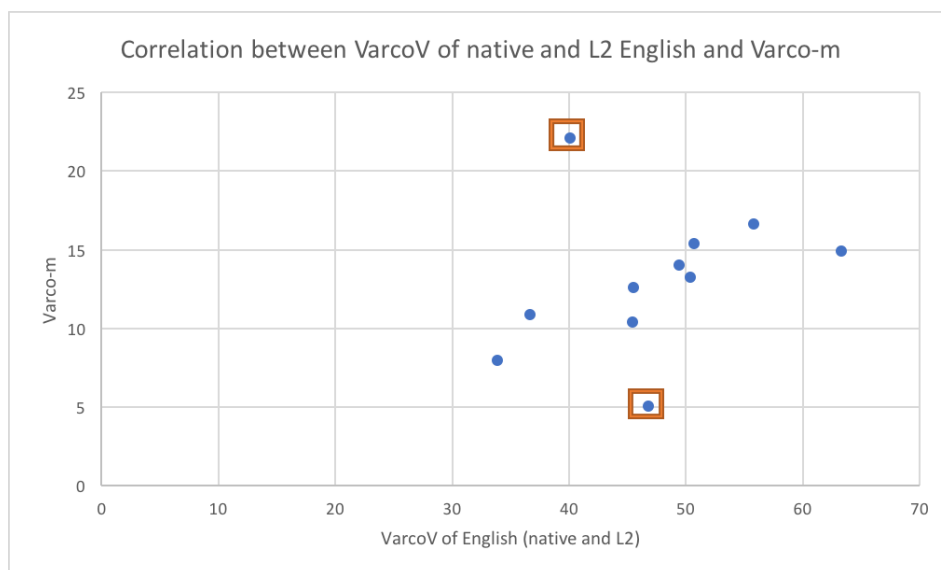


Figure 7: PMCC of L1 and L2 English VarcoV and Varco-m

Between nPVI_V of (native and L2) English and nPVI-m, there is likewise a weak linear relationship ($r = 0.21$). However, excluding Man 5 and Eng 1, PMCC results show a moderate positive relationship ($r = 0.52$).

As for correlations between L1 speech only and musical performances, there is a moderate linear relationship between the VarcoV (L1 English and L1 Mandarin) and Varco-m scores ($r = 0.4$); likewise, for the nPVI_V (L1 English and L1 Mandarin) and nPVI-m scores ($r = 0.4$). Figure 8 shows the graph of native VarcoV and Varco-m, with two outliers: Man 5 and Eng 1 once more. Again, results of the PMCC, excluding Man 5 and Eng 1, show a strong linear relationship between VarcoV and Varco-m ($r = 0.82$); for the nPVI_V and nPVI-m, results show a moderate linear correlation ($r = 0.67$). This suggests that there is some correlation between vocalic interval variability and quaver beat interval variability, at the individual level, particularly when participants Man 5 and Eng 1 are excluded.

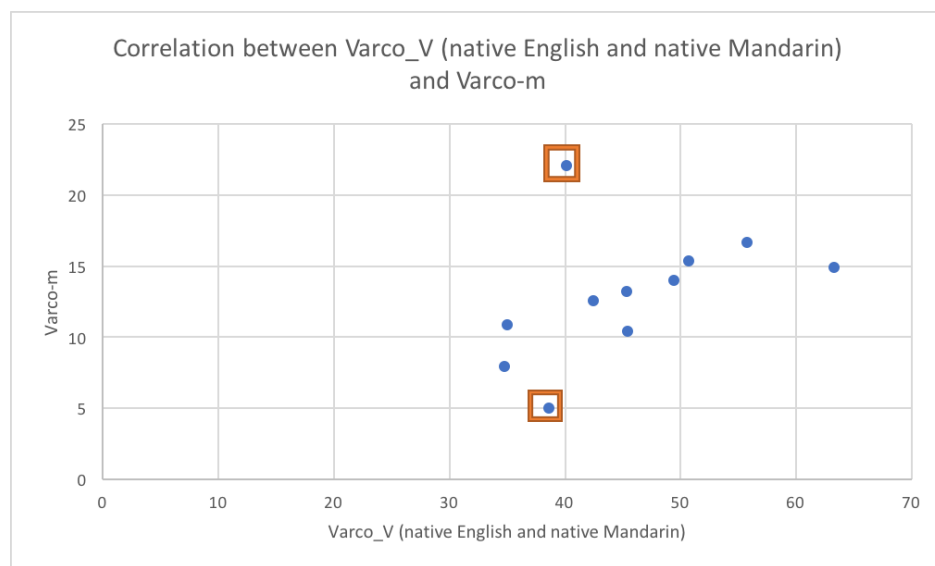


Figure 8: PMCC of VarcoV (L1 English and L1 Mandarin) and Varco-m

Interestingly, comparisons between the English (both native and L2) VarcoC and Varco-m scores of each participant produced a *strong* positive linear relationship via a firm linear rule ($r = 0.7$); see Figure 9.

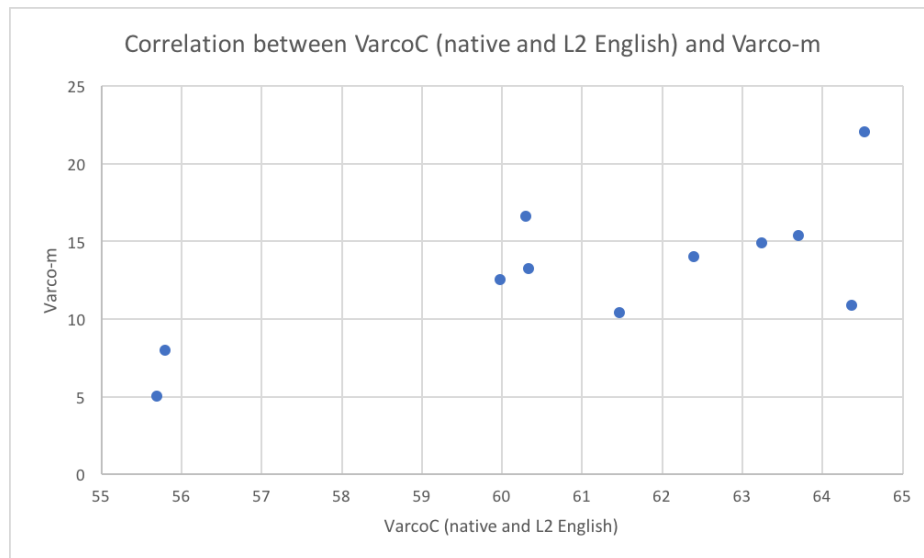


Figure 9: PMCC of English (L1 and L2) VarcoC and Varco-m

This suggests that, similarly to vocalic measures, those speakers with a higher Varco measure of consonantal interval variability also have a higher Varco measure of quaver beat interval variability. However, comparisons between English (L1 and L2) %V and Varco-m showed only a weak negative linear relationship ($r = -0.28$). These results put into question the predominant use of vocalic durations as the *sole* comparison with rhythmic measures in musical performance (although it should be noted again that in all comparable studies to date, measures of *note*, rather than *beat* durations, have been used). It is possible that, at the individual level, both vocalic and consonantal interval variability is in some way correlated with musical interval durations, provided that the musical durations are based on beat, rather than score-based (e.g. Patel and Daniele, 2003; Huron and Ollen, 2003) or performed, note-based intervals (e.g. Carpenter and Levitt, 2016; McGowan and Levitt, 2011).

Summary of Statistical Findings

Table 7 summarises the findings for statistical difference between language groups:

Table 7: Statistical difference between language groups

Language pairing	Measure	Significant difference	Higher scores
L1 and L2 English	nPVI_V	no ($U = 7$)	
"	VarcoV	no ($U = 7$)	
L1 English and L1 Mandarin	nPVI_V	no ($U = 9$)	
"	VarcoV	yes ($U = 2$)	L1 English
L1 English and L1 Mandarin (<i>without Eng 1</i>)	nPVI_V	no ($U = 7$)	
"	VarcoV	yes ($U = 0$)	L1 English
L1 and L2 English	VarcoC	no ($U = 6$)	
"	%V	no ($U = 11$)	
L1 English and L1 Mandarin	VarcoC	no ($U = 6$)	
"	%V	yes ($U = 2$)	L1 Mandarin
L2 English and L1 Mandarin	VarcoC	no ($U = 7$)	
"	%V	no ($U = 4$)	
L1 and L2 English	speech rate	yes ($U = 3$)	L1 English
L1 English and L1 Mandarin	speech rate	no ($U = 9$)	
L1 English and L1 Mandarin	overall tempo	no ($U = 12.5$)	
L1 English and L1 Mandarin	nPVI-m	yes ($U = 2$)	L1 English
L1 English and L1 Mandarin	Varco-m	yes ($U = 3$)	L1 English

Results show significant language-based grouping between L1 English and L1 Mandarin speech measures of VarcoV, both including and excluding Eng 1. There is also significant language-based grouping between L1 English and L1 Mandarin speech measures of %V. L1 English and L2 English also showed significant language-based grouping for speech rate, with L2 English having a slower rate than native English. Finally, there is significant language-based grouping between the musical performances, both for nPVI-m and Varco-m.

Table 8 summarises the findings for speaker-based correlations between speech and musical performance:

Table 8: Intraspeaker correlation between speech and musical performance

Language	Language measure	Musical measure	Correlation
L1 and L2 English	speech rate	overall tempo	weak ($r = 0.16$)
L1 Mandarin	speech rate	overall tempo	weak ($r = 0.24$)
L1 English	speech rate	overall tempo	weak ($r = -0.14$)
L1 and L2 English	VarcoV	Varco-m	weak ($r = 0.28$)
" excluding Man 5 and Eng 1	VarcoV	Varco-m	strong ($r = 0.85$)
L1 and L2 English	nPVI_V	nPVI-m	weak ($r = 0.21$)
" excluding Man 5 and Eng 1	nPVI_V	nPVI-m	moderate ($r = 0.52$)
L1 English and L1 Mandarin	VarcoV	Varco-m	moderate ($r = 0.4$)
"	nPVI_V	nPVI-m	moderate ($r = 0.4$)
" excluding Man 5 and Eng 1	VarcoV	Varco-m	strong ($r = 0.82$)
"	nPVI_V	nPVI-m	moderate ($r = 0.67$)
L1 and L2 English	VarcoC	Varco-m	strong ($r = 0.7$)
"	%V	Varco-m	weak ($r = -0.28$)

Results show strong intraspeaker correlation between the measures of VarcoV in L1 and L2 English speech and Varco-m *if Man 5 and Eng 1 are excluded*. Likewise, for the measures of VarcoV in L1 English and L1 Mandarin and Varco-m, there is strong intraspeaker correlation *if Man 5 and Eng 1 are excluded*. Furthermore, there is a strong correlation between the VarcoC measures of L1 and L2 English speech and Varco-m. Results show moderate intraspeaker correlation between the measures of nPVI_V in L1 and L2 English, *excluding Man 5 and Eng 1*, and nPVI-m. Likewise, a moderate correlation exists between both the VarcoV and nPVI_V measures in L1 English and L1 Mandarin speech and the Varco-m and nPVI-m measures. Finally, a moderate-strong correlation exists between the nPVI_V of L1 English and L1 Mandarin speech, *excluding Man 5 and Eng 1*, and nPVI-m. These findings point to an interesting correlation, in terms of durational variability, between an individual's

speech performance and her/his musical performance, irrespective, or over and above, language-background based influences. This suggests that while language-background may exert some influence over the execution of expressive nuance in musical performance, this is not the only variable of influence, and that the degree of expressive nuance that an individual shows is also mediated by a more fundamental capacity underpinning both musical and speech performance. Interestingly, however, not all participants follow this trend, with participant Eng 1 showing comparatively low durational variability in speech but very high variability in musical performance, suggesting an ability to manipulate these two types of ‘performance’ in different ways.

3.9 Post-experiment Auditory Impression Study

Following the statistical analysis of my resulting data, a follow-up auditory study was conducted with the recruitment of a professional classical pianist. The pianist was asked to evaluate the musical performances in a blind listening test, based on the amount of expressive *rubato* (i.e. timing variability) heard. The goal was to determine if the musical measures of Varco-m and nPVI-m were reflected in the expert musician’s auditory impressions of the performances. This also serves as a preliminary study for further perceptual work.

The recruited pianist was a graduate of the musical conservatoire where the performance participants were recruited. The pianist’s language background was native Polish and English L2. The same recordings used for annotation, extraction, and analysis in the main experiment were used in the listening study. The audio was played through a laptop computer, and the evaluator listened with noise-cancelling, around-ear headphones. After each performance was heard, the pianist was asked to rate the performance on a scale

of 1-10, with 1 = 'very little expressive timing variability' and 10 = 'extreme expressive timing variability'. At the request of the pianist, two further sub-categories of expressive timing variability were determined: i) local expressivity (small unit; note-by-note) and ii) global expressivity (large phrasal unit; overall line); these sub-categories were scaled 1-5. At the discretion of the evaluator, recordings were played more than once; some recordings were listened to only once, others were played again at the end of the listening session. The order of the performances was selected by the researcher and consisted of a random mix of both language groups; the evaluator was told that the performers were either native Mandarin or native English speakers but was not given any further information.

Table 9 shows the results of the auditory evaluation of the musical performances compared with the measures of musical rhythm metrics:

Table 9: scores of auditory evaluation and musical rhythm measures

Participant	Small units	Overall line	General	Varco-m	nPVI-m
Man 1	3	2	5	13.25	10.43
Man 2	2	2	4	12.57	11.76
Man 3	1	2	3	7.96	8.24
Man 4	4	2	7.5	10.89	10.14
Man 5	1	1	1	5.04	4.58
Eng 1	4	4	8.5	22.07	14.21
Eng 2	4	2	6	14.02	11.51
Eng 3	2	3	5	10.39	11.22
Eng 4	2	1	3.5	15.36	12.68
Eng 5	2	2	4	14.89	12.15
Eng 6	2	4	5.5	16.65	12.33

Statistical analysis for language-based grouping of all three categories show no significant difference (General: $U = 9$, $p = 0.32$; Small unit: $U = 11$, $p = 0.52$; Overall line: $U = 8.5$, $p = 0.27$). Analysis for intra-speaker correlation between the evaluator's general scores and Varco-m shows a strong linear correlation ($r = 0.7$); for the general scores and nPVI-m, results show a moderate-strong linear correlation ($r = 0.68$). Comparisons of small unit scores and nPVI-m shows a moderate correlation ($r = 0.5$); likewise, for overall line and Varco-m ($r = 0.6$). Interestingly, Man 4 seems to be an outlier in the correlation between impressionistic and metric scores; with Man 4 excluded, results show a stronger linear correlation between the general scores and Varco-m ($r = 0.9$) and nPVI-m ($r = 0.8$).

The lack of support for language-based grouping in this auditory impression study must be discussed: results seem to suggest that blind musical listening does not discern between language groups. This does not align with the results of the musical rhythm metric scores, which have shown language-based grouping, and would appear to throw into question the use of quantitative data to measure for language-based grouping of expressive *rubato* in musical performance; after all, judgments of performance in Western classical music are made by human listeners and not with applications and computers. However, as this was only a follow-up study using one evaluator with a particular language background, results cannot be claimed as conclusive. In addition, the quantitative analysis conducted here was on durational variability only, and did not investigate other possible parameters of musical expression, such as intensity variation and other performative factors, that are amenable to quantitative analysis. Thus, the measurements capture only part of what a human evaluator is attending to. Furthermore, although the evaluator in my study had no visual cues, real-world musical listening experiences are often not blind and the visual information of a performer is suggested to create bias in judgments of expressivity in

performance (Peynircioglu, et al, 2018). Further perceptual studies on language-based discernment are needed, using more participants, possibly also from a range of musical abilities and experience, and with careful consideration of the instructions given. Further performance parameters, such as intensity variation, are also needed to construct a more complex measurement of the data on which the auditory evaluation is based.

Although results of the impressionistic evaluation did not show language-based grouping, there was a strong indication of speaker-based correlation between the judgments and the musical metrics. This implies that although listeners are unable to discern the native language of a performer, they can discern the existence and (to a certain extent) the amount of expressive timing variation in a performance; furthermore, due to the strong correlation between a listener's impressions and the measures of Varco-m and nPVI-m, this suggests that rhythm metrics may be effective measures of expressive timing variability in musical performance.

Further study is necessary; the recruitment of more evaluators of different language backgrounds and training (e.g. other instrumentalists, non-musicians, phoneticians) would be illuminating. More rigorously controlled perceptual studies of this nature could shed light on what is considered 'expressive' to different people depending on their own musical and/or language background.

3.10 Discussion and Conclusion

To summarise, initial statistical analysis of the experimental data seems to show no support for language-based grouping of vocalic durational variability between the L1 English and L2 English speech, but it does provide support based on L1 English and L1 Mandarin measures of VarcoV and %V. As for nPVI_V, results may be different if pre-final and final phrase

syllables are included in the measurements.³³ Furthermore, in measurements of L1 English and L2 English, there seem to be outliers within the two language groups, namely Man 1 and Eng 1. Unfortunately, statistical testing for language-based grouping could not be done with the exclusion of Man 1, as the minimum sample size for calculation is 5. However, simple observation of the data without the score of Man 1 reveals that all the scores of L2 English are lower than those of L1 English speech.

It is possible that varying L2 English proficiency levels may have influenced the Mandarin scores. Li and Post (2014) suggest that differences in L2 English proficiency among L1 Mandarin speakers correlate with differences in the vocalic metrics of their L2 English speech; the vocalic metric scores (L2 English) of lower level L1 Mandarin learners of English in their study were significantly lower than higher level learners.³⁴ VarcoV and nPVI_V scores in particular seem to discriminate well in terms of proficiency levels. Although my study did not control for English proficiency levels of the Mandarin participants, a self-rated level of proficiency in understanding and speaking English was included in the participant questionnaire (see Appendix D). Of interest is Man 1, who had the highest nPVI_V and VarcoV scores in L2 English; according to the study by Li and Post (2014), Man 1 would have been likely to have a higher level of English proficiency than the participants who scored lower in L2 English. However, Man 1 stated that she/he had achieved only a 5.5 on the International English Language Testing System (IELTS), which is equivalent to B2 on the

³³ See Li and Post (2014) for analyses and comparisons of the following prosodic lengthening properties in Mandarin, German, and L1 and L2 English speech: accented syllables, phrase-final syllables, and accented phrase-final syllables.

³⁴ In Li and Post (2014), participants with L2 English levels of B1 and C1 of the Common European Framework of Reference (CEFR) were recruited.

Cambridge English Scale.³⁵ This, combined with the limited time (compared to the other participants) that Man 1 has spent in an English-speaking country, suggests that English proficiency may in this case not be a contributing factor to the differences in metric scores within Mandarin participants. Nonetheless, further investigation while controlling for proficiency levels is required.

Between L1 English and L1 Mandarin speech, initial statistical analysis also suggested no significant language-based grouping. However, as Man 1 was no longer an outlier in the L1 Mandarin speech scores, it was possible to exclude Eng 1 from the calculations (with a remaining sample of 5); interestingly, while results for nPVI_V still showed no significant difference, VarcoV measures revealed that all the scores of the L1 Mandarin speech were lower than those of the L1 English speech. This suggests that the VarcoV metric may be more effective than nPVI_V at expressing the possible influence of language-background on musical performance.

Unlike the speech measures, initial statistical testing of the musical performances showed support for language-based grouping, with greater variability of quaver beat durations in the performances of native English speakers compared with native Mandarin speakers; this would appear to support my first hypothesis, that native language experience may be influencing the musical expression in performance. Among the Varco-m scores, only Eng 3 seems to be an outlier, with a lower score than some of the Mandarin participants.

For further investigation into the outliers, namely to determine certain speaker-based differences among the participants, individual correlations between music and

³⁵ The music conservatoire attended by the participants in my study require a minimum score of 5.5 on the IELTS for admission.

speech were calculated. Initial calculations showed *only very weak* correlation between vocalic durational variability and quaver beat variability, within speakers (for both L1 English/L2 English and L1 English/L1 Mandarin speech). However, with the exclusion of outliers (this time **Man 5** and, again, Eng 1), results show a strong correlation between the speech and music of the participants. This seems to support my second hypothesis on timing variability as an element of expression that is individual and affects both speech and music. Furthermore, analysis of consonantal durations in L1 English/L2 English speech suggested significant speaker-based consistency between consonantal durational variability and quaver beat variability. This is interesting; as mentioned above: consonant measures (along with %V) usually reflect phonotactics, which may explain why there was no significant difference between the scores of L2 English and L1 English speech (as the L2 English speakers would have had to change the syllables in order to show significant difference). Perhaps the expressive rendering of a score through performance correlates with certain structural elements within an individual's rendering of text through read speech. This suggests that, to test for correlation between read speech and performance of composed music, other variables, beside vocalic durations, may be useful.

A discussion of the various outliers, discovered during statistical analysis, is necessary. The first outlier, Man 1, is shown to have noticeably higher durational variability in both speech (native Mandarin and L2 English) and music, particularly for her/his language group. This suggests that Man 1 exhibits a high amount of timing variability in either context. The second outlier, Eng 1, is shown to have unexpectedly low durational variability in speech but the highest variability out of all the participants in musical performance. This suggests that Eng 1 utilises timing variability as an expressive device in music to a higher extent, but the opposite applies to her/his speech. Another outlier, Man 5, is shown to have

the lowest durational variability in music but has relatively high variability in speech, particularly L2 English. This suggests that Man 5 exhibits a high amount of timing variability in speech, yet exhibits a low amount in music.

As for the remaining participants, the following observations may be formed: i) Man 3 exhibited equally little durational variability in both speech and music, ii) Man 4 and Man 5 show an inverse of the expected relationship between VarcoC and %V in their native Mandarin speech (for reasons yet unknown), iii) Man 4 exhibited less durational variability in L2 English speech than in music, and iv) Eng 3 has the lowest VarcoV and Varco-m of those in the native English group.

The results of my study seem to suggest that native language may indeed influence the utilisation of durational variability within one's expressive performance to a certain extent, but also that timing variability as an element of expression is highly individual and may affect both speech and music. Furthermore, as the data from certain participants demonstrate, timing variability may not necessarily be expressed equally in speech and music; some seem to show more variability in one than the other, some in neither, and some in both. The reasons for this are unknown; further investigation is necessary.

Auditory impressions of the timing *rubato* of the musical performances revealed a distinction between performances with the low variability, compared to those with the high variability. Comparison between the ranking and the metric scores showed significant correlation; notably, the two extremes of lowest variability and highest variability were consistent in both measures. Specifically, Eng 1 scored the highest variability in both the rhythm metrics and the impressionistic ranking. Man 3 and Man 5 scored the two lowest variability in the rhythm metrics, as well as the two lowest rankings. Also, there were confounding differences in the way *rubato* was used by the participants; while the majority

used timing to show groupings of two-bar units, Man 2 showed variation within smaller units (per musical half-bar) and Man 5 showed almost no timing variation. Eng 2 showed both two-bar as well as half-bar units, creating a hierarchy of timing variation. The majority showed significant lengthening at the highest point of prominence (the climax) of the entire excerpt; however, Man 3 and 5 showed more strengthening of dynamic (getting louder) than lengthening. The majority of participants exaggerated the length of longer note values (i.e. the dotted quavers, particularly in bars one and three), which are present at the beginning of each two-bar unit; however, Eng 3 also showed a tendency to lengthen the ending of each unit (i.e. the end of bars two and four).

The variety of differences in expressive *rubato* as well as the varying impressionistic rankings of the participants suggest that expressive timing variation is highly individual, over and above any language group differences. However, the rhythm metrics of the musical performances suggest that native Mandarin speakers perform with less timing variation than native English speakers. To formulate a possible explanation for this, I now refer to the multi-componential model of speech rhythm by Post and Payne (2018). This model functions under a holistic theory in which the ‘experience’ of rhythm in speech (which involves both the perception and production of rhythm), as well as in other cultural constructs such as music, dance, and visual arts, is posited as an experience of the ‘temporal [or spatial] organization of a sequence of similar events or objects, and of other parameters defining the relationship between these events/objects’ (Post and Payne, 2018: 7). This theory is particularly useful due to its recognition of the parallels between speech and music, and its belief in a common neurobiological and ecological grounding for experiencing rhythm in both domains. According to Post and Payne, speech rhythm is both multi-systemic and multi-parametric. In other words, the experience of speech rhythm involves multiple,

integrated systems that are both structural (including lexical and phrase-level prominence, phrase boundaries, and lexically contrastive length) and performance-based (speaker idiosyncrasies, affect, speech rate, and speech style). Speech rhythm also involves different phonetic parameters in which the systems are signalled, namely duration, intensity, pitch/fundamental frequency, and spectral variation. To understand speech rhythm as a complex, integrated, and multi-componential experience explains the varying and at times contradictory data found in research on different and dissimilar languages. This model suggests that cross-linguistic variation in the speech rhythm percept cannot be categorised by singular parameters such as duration, and that an overall impression of the speech rhythm for a certain language is generated by multiple structural and performance-based properties that are transmitted phonetically in various ways. The results, albeit inconclusive, from my own rhythm investigation would appear to fall in line with the above model; and my exploration of timing variability in musical performance leads me to extend this model to music.

It can be readily appreciated that in both speech and music, performance-based deviations from the commonalities and constraints of either a particular language or notated score are intricately woven within the experience of rhythm. Along these lines, as well as those listed above, it may be possible to use the multi-componential speech rhythm model to view musical expression in performance as a holistic phenomenon, also involving multiple systems (the prescribed notation) and parameters (the components of performance). As discussed in Chapter Two, the work of Palmer and Hutchins (2006) described certain ‘prosodic cues’ in musical performance that seem akin to speech systems;

these include both ‘preference rules’ (named ‘preference’³⁶ because there may, in any given piece, be more than one appropriate ‘musical prosody’) that delineate structure within a musical composition through segmentation (i.e. boundaries, final-phrase lengthening) and prominence (i.e. stress, accent); and performer-specific cues that emphasise particular structural and emotional intentions, depending on the interpretation of the performer (Palmer and Hutchins, 2006: 28). The acoustic parameters available in expressive musical performance are largely the same as those in speech, namely duration, amplitude, pitch, and timbre (harmonic spectrum). Returning to my results of the musical performance investigation, the idea that there are multiple systems that involve multiple parameters within expressive musical performance could be used to explain the difference in the amount of durational variability utilised between the two groups of participants. It is possible that the performers in the native Mandarin group used parameters other than duration, e.g. intensity, perhaps to a higher degree than those in the native English group, to show variability within their expressive performances. This aligns with the observation that Man 3 and 5, who displayed the least amount of durational variability both in the rhythm metrics and impressionistically, seemed to have a greater increase in amplitude at emphatic moments within the performance.³⁷ Further comparative, cross-domain and cross-linguistic studies on other parameters of rhythm, such as intensity (see He, 2012 for L1 and L2 English) and fundamental frequency (see Keating and Kuo, 2012 for English and Mandarin speech) are needed. Furthermore, an investigation of specific structural systems

³⁶ The use of the word ‘preference’ is first mentioned in Lerdahl and Jackendoff, 1983)

³⁷ This is not to say that the native-English speakers’ use of duration is exclusive; but perhaps there are differences in how the various parameters are weighted, between language groups.

of speech rhythm, such as stress (both lexical and phrase-level) and phrase-boundaries, and phrase-final lengthening, could have informative results; these systems could be studied comparatively with expressive emphasis (beat-based or note-based) and phrase-final lengthening in music to measure for amplitude or timbre variation between language groups. Additionally, while not available in piano performance, pitch variation in non-keyboard instrumental music may also prove fruitful in comparison with F0 variation in speech.

To conclude this chapter, I have found the speech rhythm model offered by Post and Payne (2018) to be a useful and relevant framework for modelling the multiple variables and complex nature of musical expression in performance, and one that can provide a foundation for further work within musical performance research, linguistics, and comparative studies between both fields. The next chapter details the second stage of my investigation and will provide a personal reflection on how these results and considerations have affected and influenced me as a musical performer; it also proposes how they might influence the Western classical music community at large.

Due to a shift in research stage, and thus methodology, the narrative style of the remaining chapters differs significantly from what has been presented so far. However, I believe it was necessary to present this juxtaposition of two seemingly disparate perspectives, in order to demonstrate the challenge I faced as an artistic researcher, seeking to unite my empirical investigation of the performance of others with my own musical practice.

Stage Two

Chapter Four: Autoethnography

4.1 Summer, 2018

I sat in the large common room of an old house, located within a hamlet that's nestled on the slopes of the Mount Canigou Massif in southern France. Through the windows, I could see the forested curves of the Pyrenees, a sight that evokes awe and gratitude in anyone who happens to take a glance. I had been invited by a friend and colleague to spend a few weeks at his music festival, practising, rehearsing, and performing chamber music. A few days into the experience, aided by the gorgeous scenery and warm hospitality, I felt compelled to reflect on how I had reached this point in my musical journey. I was the only East Asian musician at the festival, which was meaningful in that I was surrounded by cultures and languages previously unfamiliar to me. The speech of Italian, French, Spanish, German, and various accents of British English flowed around me, along with the music of Mendelssohn, Shostakovich, Enescu, Mozart, and Schubert. The sounds of the languages instilled as much wonderment and inspiration as the music, and I felt at once fortunate and grateful to be given this opportunity.

It was the day after a tiring yet gratifying performance of two great works for piano trio: Mendelssohn Piano Trio in C minor, op. 66, and Shostakovich Piano Trio, no. 2 in E minor. I've always felt more at ease in chamber music than in solo playing. Perhaps it was because, in the first case, the communicative element of performance extends not only to the audience, but also to the companions around me. When I perform chamber music, I don't have time to consider how my musical voice is perceived; I am taken in by the constant conversation and collaboration with others. Performing solo repertoire, on the other hand, is

an entirely different experience. My doubts and insecurities show themselves when I am left to communicate the music alone. When the silence surrounds me, the negative judgments come from within, criticising my technique, my interpretative choices, even my stage presence. I wonder about what other musicians think of my performance, whether my expressive decisions are suitable for the music. Because of these thoughts, my solo performances have often been the source of stress and anxiety, unlike my chamber performances, which usually bring pleasure and excitement. 'This is wrong,' I thought, 'I shouldn't have a complex over solo repertoire; how will I sustain a performing career?' Just then, someone called out to me through the window, breaking my reverie. I walked over and leaned against the ledge, feeling a slight groan of the old wooden frame as my weight settled upon it. My friends were swimming in a natural pool, situated almost precariously at the edge of the elevated property. Smiling at their waving arms, I decided to take a break from my reflections and join them.

4.2 Introduction

I effectively 'discovered' the method of autoethnography—as a way of interpreting, analysing, and reflecting upon my research—much later than when I had first heard about it. I was introduced to the method during a doctoral training session, and remember feeling rather wary towards the kind of work that was being presented as qualitative research. As I read through some examples, I couldn't help feeling critical towards what I saw then as creative liberties that some of the authors had taken in work that was supposed to be scholarly; I later discovered that I was not alone in my reaction. Autoethnography is a controversial method: first developed in the social sciences, it is difficult to define and encompasses a variety of approaches that in turn have attracted numerous different labels

(Wall, 2016). These labels, which reflect the many disciplines that have influenced its development as a methodology over time (Chang, 2008), include autobiographical ethnology, critical autobiography, narrative ethnography, reflexive ethnography (Ellis and Bochner, 2000), and even heuristic inquiry (Wall, 2006). Regardless of precise terminology, autoethnography (as coined by Hayano, 1979) has three core elements: the self ('auto'), culture ('ethno'), and research process ('graphy'). These elements may be emphasised differently, and examined in varying degrees of complexity, in different research. Some researchers stress the importance of utilising established, ethnographic research approaches to emphasise particularly the elements of 'ethno' and 'graphy' within their work (Anderson, 2006; Chang, 2008). This tends to involve field data collection through means of participation, observation, and interview; data analysis/interpretation; and report writing. For such approaches, cultural understanding of self is achieved only through a thorough examination of the cultural context. On the other end of the spectrum lies the 'evocative' tradition of autoethnography, which emphasises the importance of personal ('auto') experience. Ellis and Bochner describes this approach as the following:

I start with my personal life. I pay attention to my physical feelings, thoughts, and emotions. I use what I call systematic, sociological introspection and emotional recall to try to understand an experience I've lived through. Then I write my experience as a story. By exploring a particular life, I hope to understand a way of life (Ellis and Bochner, 2000: 737).

I realised that the work that I had been introduced to during my doctoral training was mainly of the evocative tradition, hence the use of first person voice, along with creative-writing styles of storytelling. The accessibility and readability of these text made me

rather uncomfortable due to my familiarity with research from a positivist perspective, which stresses quantitative, experimental, and 'objective' data, delivered in dry, third-person voice, as the only legitimate goal of intellectual inquiry. On the other hand, evocative autoethnography functions under the postmodern philosophy that there are many legitimate ways of knowing and obtaining knowledge, including the involvement of the self and subjectivity, and none should be privileged over any other (Wall, 2006). In this sense, evocative writing purposefully uses stories, poetry, and other literary styles to invite readers to participate emotionally and in dialogue with the author; the writing 'activate[s] subjectivity and compel[s] emotional response', and 'offers lessons for further conversation rather than undebatable conclusions' (Ellis and Bochner, 2000: 744).

By the time that I turned to autoethnography as a viable option for my research, I was much more aware of the reasons for, and benefits of, using such a method. Within social science, autoethnography has been established as a useful and accessible approach to studying difficult topics, ranging from the highly personal and sensitive (such as death, abuse, illness, and tragedy) to the less sensitive but equally relevant topics of identity, multiculturalism, familial relationships, and professional practices (Chang, 2008; Denshire, 2013). Chang lists the benefits of the autoethnographic method as the following: i) 'It offers a research method friendly to researchers and readers'; ii) 'It enhances cultural understanding of self and others'; and iii) 'It has a potential to transform self and others' (Chang, 2008: 52). I was particularly intrigued by the capacity of the self-reflexive process of autoethnography to strengthen awareness of a sense of self, as well as examining existing preconceptions and judgments (made not just by oneself but also by others) from cultural and societal standpoints. The further that my research extended toward issues related to

my own sense of identity, as well as my particular cultural position within the world of classical music, the more I gravitated towards such an approach.

Another reason for trying a different methodological approach transpired after the completion of my quantitative experiment and statistical analysis of the results. I reached the stage of interpreting my experimental data, and was faced with numbers and levels of significance that seemed to float in mid-air, suspended and lost, without a purpose. The quantitative approach had served me throughout the process of data collection and analysis, but now I had the task of interpreting the results from the perspective of a performing classical musician, and I faltered. I became aware—to an alarming extent—of the gaping distance between my empirical data, which seemed to point to language-based differences between the performances of native Mandarin and native English pianists, and my particular position as a Chinese-born, American-raised, London-residing classical pianist. I felt a strong sense of responsibility towards myself and to the entire community of performing classical musicians, especially those from non-Western cultural backgrounds, to disseminate my results in an informed, convincing, and relevant way. The contents of Chapter Five resulted as one way of reacting toward my empirical data; I sought to be more aware of the social and cultural situations and contexts of East Asian classical musicians. I delved into cultural texts on the East Asian experience within Western classical music, and was met with a discourse that seemed at the same time new and familiar to me, one from the perspective of the Asian musicians themselves. This added a further layer of reference to my work, and I needed a method that could address and make sense of the many complexities that had developed. Ultimately, I looked to autoethnography as the most suitable approach to examining my research as it evolved into a further stage, one involving elements of my person, my (mixed) culture, and my practice.

This current chapter is the written expression of my thoughts, and the emotions that appeared as I reflected upon my experimental results, my musical experiences, and my practice while in the course of informing myself through cultural texts on East Asian musicians in Western classical music. The non-scientific style and narrative approach reflect a certain transformation within myself and my work, as I had undergone a research journey that deepened my understanding and awareness of the complexity behind my initial research questions. My literary voice developed from one that functions within standard academic style to one of a more evocative nature. Each of the sections in this chapter begins with a narrative that sets the scene of an experience and reveals internal reflections upon said experience. This is followed by a dialogue through which my ideas are established and confirmed; the first voice of the dialogue, known as ‘One’, serves as the sceptic and poses concerns that the second voice, ‘Two’, attempts to answer. From Ellis (1991), I borrowed the methods of emotional recall, as well as introspection as a ‘systematic, sociological technique’ to thoroughly examine my own identity as a performing musician. Looking inward, rather than to existing published sources, I was eventually confronted by my own understanding of musical expression, which seems to have been shaped by certain traditional, predominantly Western discourses. These discourses, stemming from a singular (Western) perspective, seem to be responsible for the tensions and doubts that I’ve felt throughout my musical life.

Dialogue 1: Autoethnographic Method

***One:** I see that you have attempted to justify your use of an autoethnographic method, but I’m afraid that I’m still not quite convinced. It is easy to state that, suddenly, you*

arrived at a methodology that seemingly rescued your research from stalling at an unfortunate period, but on what basis? An intuitive notion? You can't choose, on a whim, to switch to an entirely contrasting method from the one that you had chosen from the beginning; remember that your first method directly correlates to your research questions on the speech and musical expression of classical performers!

Two: I admit that the direction of my research has drastically shifted since the start of my endeavour, but as often is the case with work on complex, interdisciplinary topics such as the relationship between language and music, the process rarely follows a single consistent and rigid method from beginning to end. Particularly in my case, as both a researcher and an active performer, the perspective that I have brought to this topic is unusual, if not unique, in the literature. I have not found any work on the language/music relationship done by performer-researchers, and the issue is further complicated because, by being a performer-researcher, I am inevitably involved as an insider. This means that my knowledge, experiences, and judgments engaged with and influenced the research that I conducted. I thought that, if that were the case, why not go all the way and have the *personal* be the focus?

One: *But what of your original research questions, which were based on quantitative interrogation? Also, your hypotheses were essentially proven by the experimental data...*

Two: Yes, exactly! I had reached so-called 'conclusions', and perhaps the researcher in me would have been satisfied with that, but the performer in me was not. I began to have more questions, like, 'Now what?', 'What do I do with this information?', and 'What

does this mean for my own music-making, as well as that of other East Asian classical performers?’ And it was then that I realised that I had to probe further than what my original research questions had managed to do; I had to confront more personal concerns, such as understanding why I was so fixated on the idea of difference in speech and musical expression in the first place. My attention shifted from a very specific and narrow focal point—on comparisons between vocalic interval durations and quaver beat durations—to broader concerns about the various experiences, discourses, and norms within Western classical music that may have led me to have feelings of difference and doubt. These concerns are simultaneously personal and broadly cultural and social, and the method of autoethnography deals exactly with such complex, multi-layered themes.

One: *Okay, but you haven’t really supported your use of the method with an awareness of other work in the literature. For instance, in the field of musical performance research, are there others who have used the autoethnographic method in relation to their own performance? I accept that there have been articles within other disciplines, particularly in the social sciences, on topics such as work activities and experiences, illness and injury, academic life, family life, and membership of alternative cultural communities (see Wall, 2016). But what about within music? I ask also because I don’t see how you can use the method to address your own musical practice, which has thus far been peripheral in your investigation.*

Two: I will address these concerns one by one. Firstly, self-reflexivity has been gaining popularity among musicians who want to examine, understand, and communicate their artistic processes and creations in relation to both personal and cultural experiences. The

introductory article by Bartleet and Ellis (2009) in their collection of music autoethnographies provides a good context for understanding the movement towards personal exploration within the field of music. If you're looking for specific examples, I can cite Bartleet (2009) as one of the first music autoethnographies that I found to be both interesting as a performing musician and encouraging as a researcher. Bartleet's work showed me that one's personal experiences and feelings during practice and performance can lead to further understanding of oneself, as well as of the cultural landscape in which one functions. In terms of the specifics of practice, music autoethnographies can be done in significantly varied ways; the collection of work in Barleet and Ellis (2009) offers contrasting approaches and processes...

One: *Yes, I've looked at the collection, and honestly, only some studies seem directly relevant to musical performance within the Western classical canon.*

Two: I agree with that; I found Stephen Emmerson's (2009) chapter to be more familiar and practical than some of the others, perhaps simply because he is a classical pianist and wrote about how his interpretation of a solidly canonic piece for voice and piano was informed by his practice.

One: *I didn't really see how his work dealt with larger cultural, social, or political considerations...*

Two: I think it was quite specific in its focus, and thus its target audience; he described how rehearsing with different musicians and instrumentalists had influenced and altered his interpretation of the music, so perhaps there is a social element to his work.

One: *Is that what you are going to write about, later?*

Two: There are moments in this chapter that are dedicated to the discussion of specific details within the music that I have worked on, and as a significant portion of my repertoire has been chamber music, there will be mention of my work with other musicians. I will mainly discuss notable concepts and practice methods gathered from my work with established instrumentalists and faculty in chamber music masterclasses. These methods target specific sections within the music and are informed by certain characteristics in the speech of different languages.

At this point, I'd like to reference Doğantan-Dack (2012a) as an example of autoethnography in relation to artistic practice within a chamber music format. Using live performance as the research focus, Doğantan-Dack describes specific details and situations during rehearsals and live performances to show a unique process of gathering knowledge. Furthermore, rather than having solely the practice and rehearsals inform the live performance, she also believes that the live performance can inform and illuminate the practice by serving as the research material. She also encourages performers to 'reflect upon the nature of the reflective component itself, contemplate the most appropriate means of documenting the event, and present the thought processes surrounding the performance' (Doğantan-Dack, 2012a: 40); autoethnography, she believes, is a highly appropriate method for such goals.

One: *And how does that create meaning in the cultural, social, or political dimension?*

Two: Doğantan-Dack argues that giving agency to the performer in research of this kind is a political act; too often, the role and status of performers are not in balance with those of the researcher, to the point where the performers' voices are anonymised. It comes down to giving voice to performers within the dominant disciplinary discourse. Also, by focusing on chamber music, she highlights the social dynamics among co-performers within an ensemble as a significant source of knowledge. Although my research does not offer such a detailed view of the rehearsal dynamics of my own ensembles, my experiences and interactions with others have undeniably impacted my views on musical expression, communication (both verbal and musical), and performance.

4.3 Practice and Musical Experiences: France and Belgium

It was back, the pressure to be 'authentic', whatever that even means. I gave a big sigh and slumped downwards on the piano bench. 'Negativity is not good for your posture,' I thought and straightened my spine. Staring at the score of Schubert's Impromptus, op. 142, I could feel a tickling in my nose and the sour approach of tears. Sniffling, I shook my head and sought a more positive and calm voice within. This voice, perhaps it's my so-called instinct, was suggesting that I approach these pieces as if for the first time, using knowledge gained from the new experiences and discoveries that had occurred since the last time I confronted this music. The voice told me to shed my old skin, old habits, and old traumas; 'Work away from the piano', it whispered. Obeying, I picked up the score and walked to my bedroom.

In my room, which was just steps away from the common room and the Bechstein grand that had been my companion for several weeks, I opened the score to the first page. Running my hands over the paper, I suddenly thought that the source of this music could not be found in my limbs. My hands produced nothing but a swooshing sound as I slid them over the page. The music of Schubert exists within the score, and even then, incompletely so. There was only so much that could be represented by standard notation; by text and symbols. The 'real music' surely existed in a different plane, accessible to all who have put in the work, the time, have absorbed a wealth of influential stimuli, and have achieved ease and agility at their instruments...

Lying on my bed, I began to think about my earlier conversation with an established Baroque cellist who happened to be in residence at the festival. I had expressed my concerns about my own musical voice and whether I had the authority to express Western classical music. His response was to describe an experience that had occurred while he was living and learning from a master musician in India. In his distinctively Italian-accented English, he told me of an evening when he was listening to a performance of his teacher and other distinguished musicians. It was very late, as the concert only began around midnight, and the audience had gathered outside, sitting against the trunks of mango trees or directly underneath the stars. The atmosphere was electric, and the cellist became absorbed with counting the beats of the talas, finding himself entrapped within the cycles and not being able to break free. Suddenly, as he felt his mind and body completely possessed by the rhythm of the music, a string on his master's sitar snapped. The other musicians stopped playing and, completely unfazed, his master began putting on a new string. The cellist told me that in those moments of silence, he was still unable to get out of the cycle of counting. He was locked within the rhythm, unable to move or make sense of his surroundings. After

what seemed like an eternity of counting in silence, his master finished tuning and started to play again. What shocked the cellist to the core of his being was that the music began exactly on the first beat of the tala that had been continuing in his mind.

After the concert, the cellist rushed to his master to describe what had happened. His master simply smiled and responded that what he had experienced was the true essence of music, which exists beyond any worldly and human intervention. It exists, absolutely and completely, in what can be thought of as another dimension. All performers who have musical knowledge and have mastered the craft of producing music, whether on an instrument or through the voice, are able to access this essence. And the 'voice' of a musical performer is simply a vessel in which the music can be made to sound. As I struggled to fight back the tears that appeared at listening to such a magical story, I was reminded of something that my piano teacher in London once said. I couldn't recall what piece I was playing, but remembered that it was at the very opening that my teacher stopped me and said, 'Music exists as itself, it is not something you mould into existence with your hands. Trust that the music is there, and simply let it sound.'

I knew what all of this was meant to do, it was meant to free me from the self-conscious, critical voice in my head that judges and doubts everything that I play and makes me appear closed, or careful, or even 'trying too hard' on stage. If I could believe that music exists as itself, that anyone with enough skill, knowledge, passion, and experience could access its essence, then I could overcome my difficulties and be a confident, convincing performer. The concept of universality in music – about which I'd had so many doubts – appeared yet again in front me, and all I had to do was wrap myself up in its comfort. But that didn't happen. I couldn't throw off the nagging doubts within.

I spent the last week of the festival revisiting my solo repertoire. I knew that when I returned to London, I'd have two weeks to prepare before I left for Belgium. There, I'd be playing Beethoven's last piano sonata, op. 111, as well as the Schubert Impromptus, for two highly regarded pianists. One of them was a musical idol of mine. I wanted to prove myself, to show my strengths and gain the approval of these masters. As I practised over the weeks, I thought about the musical voices of pianists whom I respected and admired: Alfred Brendel, Richard Goode, Andras Schiff, among others. For Beethoven and Schubert, these were the voices that I'd listen to, that seemed 'right' to my instinctive tastes. Brendel was particularly influential; when I listened to his playing of the Schubert Impromptus, I felt that there could not be any other way of expressing this music. As I practised the same pieces, I felt an overwhelming urge to judge my playing in comparison to that of Brendel. I was aware that it was a foolish urge, one that was not only unhelpful but likely to prove detrimental to my progress: a source of negativity, depleting my energy and drive. I couldn't help it, though; it was a habit I had formed in my youth, and although I could usually stop myself from relying on recorded interpretations for reference, I was always left with feelings of inadequacy. For a few moments, I sat in a 'pool' of negative emotions, then shook myself and returned to my practice.

Towards the end of the festival, I gave an informal performance of the third Impromptu, in B flat Major, to a mixed audience of musicians and non-musicians. I knew I was at an early stage in my revision work on it, and tried to focus on expressing the music in what I thought of as an intermediate state. Rather than seeking to achieve the ultimate in quality of performance, I allowed my feelings and sensations to govern as I played. A brief portion of the performance was recorded, and as I listened back, I noticed that my musical voice seemed slower and more expansive than usual. I decided that this was due to a

combination of letting my emotions govern the performance, as well as not having enough time with the piece, or confidence, for it to flow. The feedback I received was generally positive, with mentions of my beautiful sound and atmosphere; however, one of my friends, a violinist, thought that I could 'let go more', particularly during a section of intensity and impassioned struggle. I was surprised at her comment because, during the performance, I felt that I had played with proper character and intensity. Her response was that it still needed more. I wondered whether this was due to my lack of adequate preparation, or simply that my musical voice was incapable of expressing extreme emotions. The phrase of the world-renowned musician, quoted in Chapter One, suddenly came back to me: 'It is difficult for Asian people to show extreme emotions in public'. I was reminded of my frustration at being generalised and misunderstood.

After the festival had ended, I returned to London and practised intensively for two weeks. I felt my physical strength, technical agility, and musical conviction reach a suitable state. When I arrived in Belgium, I felt prepared both mentally and physically. My masterclass experiences with both teachers were positive and uplifting; I was praised for my technical ability and having 'strong fingers', as well as for my musicality and ability to 'listen' while I played. I was very pleased with this outcome, as I was distinctly aware and diligent in my practice of these elements. The masterclasses, themselves, were interesting to reflect upon. I became aware of the differences in musical voice, not only of the students but of the teachers as well. I found myself sensitive to differences in performance style and expressivity, no doubt influenced by geographical, educational, and, perhaps, language backgrounds. It was particularly fascinating to witness the faculty rehearsals and performances of chamber music; in these situations, the communication and interaction between musicians with contrasting voices could be seen on full display. There were, as was to be expected,

disagreements on which compromises had to be made. Everyone was very professional, but I could sense a certain discomfort when voices clashed. This intrigued me; it was highly plausible that the musical voices of experienced and distinguished performers should be more present and strongly established than that of students and/or younger performers. Another interesting observation related to the contrast of teaching approach within the piano faculty; as the teachers changed between weeks, I could only prepare a single programme for both. I was not without experience of receiving tuition from different teachers on the same piece (I had, for several years, been in shared studios at both of my conservatoires). However, it was still quite an experience to receive two sets of such intensive instruction in the span of two weeks; the similarities and differences between the two teachers were highlighted to a greater degree. Fortunately for me, these experiences led to useful observations for my research. For example, a noticeable difference in instruction was prompted by the opening of Beethoven, op. 111: the teacher in the first week wanted significant space between the chords, telling me to wait until the moment that felt even too late to play the next chord; the teacher in the second week preferred a much faster and urgent tempo; when I played the opening at the first speed, she commented that it was not 'in tempo'. Afterwards, I reflected upon the teachers' contrasting speech rates when they spoke in the lessons: the first teacher had a slow, drawn-out pace while the second teacher had a lyrical, impulsive element. This seemed to be reflected in their own playing, as well; their respective performances of the first and second piano trios of Mendelssohn displayed obvious differences of pacing. My thoughts, at the end of the course, related to the subjectivity of musical pedagogy; there are, of course, common instructions that relate to structural and stylistic concerns shared across Western classical music. However, this doesn't preclude the existence of idiosyncrasies, differences in 'taste', differences in 'schools' of

musical practice, and much more, all of which are extremely complex and difficult to classify. Inevitably, the development of a musician's own musical voice, as well as the confidence to realise that voice, becomes of central importance. 'This is what will transform a student into a professional musician and allow her or him to sustain a thriving musical career', I resolved.

I breathed a long sigh and kicked off my concert shoes; black, heeled things with a sneaky platform that made them look higher than they really were. I was in the ladies' toilet, changing out of my concert dress and into a much-more-comfortable attire of jeans and a red jumper. Smiling slightly, I thought through the events that had just occurred. I performed two of the four Schubert Impromptus in a student concert. It was a short programme, but as it was my first performance at the festival in Belgium, as well as my only solo performance to a large audience in several months, I was pleased with the outcome. Hours before the concert, I was alone in a practice room, preparing both mentally and physically to get into the 'zone' of performance. Seeking an inspired mind-set, I instinctively searched online for a video of a pianist whom I had met once and whose playing I greatly admired. I had watched her videos before, and would feel an instant connection to her expressivity and her stage presence. She seemed, to me, entirely committed to the music; it was as if her playing 'spoke' to me directly. This direct quality of communication was something that I admired most in certain performers; I'd feel uplifted and connected, both emotionally and musically, whenever I sensed that a performer is 'speaking' the music.

Lying down on several practice room chairs that I made into a makeshift bed, I watched her play the Bach-Busoni Chaconne. I knew that the concert was just in a few minutes, and in a few minutes, I'd have to pack up my clothes and scores and walk to the hall. However, I became engrossed by the energy and conviction of the pianist's music, and

could feel her power surge through my own body. I realised, in that moment, 'It doesn't matter how you say the music, you've worked that out already. Now is the time to just speak it!' I knew that doubting one's musical decisions during a performance was more disastrous than making a memory slip or failing at a technical run. It would show a lack of conviction, which in turn weakens the connection with the music for the performer, as well as the enjoyment for the listener. As I arrived and waited backstage, I forced the doubts and fears from my head, feeling only the rush and excitement of stepping up to the piano underneath the blazing lights. As I took my bow and sat at the keyboard, I gave myself only one purpose: to speak. And the concert was a success; the audience's enjoyment and appreciation were palpable. I also received praise from fellow students, teachers, as well as the artistic director of the festival. That night, I made a mental note to remember this moment and the lesson it offered me, that no matter what insecurities I may have about how I played, the performance state of mind takes precedence in those crucial moments onstage. 'If I have prepared to the best of my ability, I must then simply allow my voice to speak. Anything else would be a disservice, both to the music and to myself', I decided. However, in the back of my mind, I knew it would not be this easy, not every time. Each performance, each event would have a different effect and outcome. I also felt that I still preferred the satisfaction of having experienced music onstage with others; that the exhilaration of playing my own part, while having to react, respond, and adjust to other voices at the same time gave a greater sense of joy and accomplishment. The looks of gratitude and fulfilment on the faces of my companions after a successful chamber music concert was something that I'd miss whenever I performed alone. 'In fact, you don't really even like the spotlight', said a voice within, 'you've always had to be inspired by someone else, even become someone else in order to

make it through....' Shaking those thoughts away, I left the toilet and distracted myself with plans for dinner.

Dialogue 2: Chamber Music

One: *You made several remarks in your series of recollections that were never elucidated; the first of which seems quite serious and deserving of further unpacking. It is related to your idea of a 'performing career' and your unwillingness to accept what seems to be a predilection for ensemble-playing.*

Two: You have pointed out a major issue that has been revealed to me through self-reflection. Since my first experiences of learning chamber music as a first-year undergraduate student in Boston, I have struggled to come to terms with my enjoyment and satisfaction at devoting time and energy towards something that was not solo repertoire. Although chamber music is certainly considered a worthy artistic pursuit, one that fosters further knowledge, growth, and strengthening of both musical and personal relationships, there seems to be an unspoken acknowledgement that pianists who pursue 'collaborative' career-paths do so because they found they could not cope with those of a soloist. Chamber music seems to be a kind of elevated extracurricular activity, a notion supported by comments like, 'even the most successful soloists play some chamber music', as well as by the ubiquitous presence of a chamber music piece at the semi-final (and not final) stage of many prestigious international piano competitions (but not all: the Tchaikovsky International Piano Competition being a notable example). Pursued as a side-line activity in a classical pianist's career, it is a sign of high-level, well-

rounded musicianship; chosen as the focus, however, it seems the 'less than ideal', 'second choice' to the primary goal of becoming a soloist.

One: *I must interrupt you here; are you claiming to speak for the entire classical music community? And if so, what evidence can you provide for making such claims? Surely not every pianist trained in the Western classical tradition shares your perceptions on chamber music?*

Two: What I have described stems purely from my own views, formed, on the one hand, through observation and, on the other, through opinions both explicitly and implicitly communicated by fellow musicians, teachers, friends, and even my own parents, whose views on chamber music became more positive as I gained more success through ensemble work. My words here represent only the view of chamber music's value that I myself have formed, but this is valuable in itself, as my prejudices have been cultivated by and within the Western classical music community. Furthermore, and more importantly, I am using this opportunity of self-reflection quite explicitly to confront my own prejudices towards chamber music and acknowledge the transformation I can feel is taking place.

One: *And what is this transformation?*

Two: Reflecting upon the past few years of my research and musical experiences, I have recognised that my engagement in chamber music has contributed significantly to my study on the speech and music relationship, as well as to my musical development.

Conversation with other musical voices, inspiration from sources outside of my own instrument community, and engaging in learning and discussion with a variety of peers and faculty have contributed to both my musical and personal growth. My prejudices, which previously were hidden by an active engagement in chamber repertoire, were made clear to me: I was not willing to focus on the art of ensemble-playing. This was a revelation, as I had always thought of myself as different from those with rather narrow-minded views about performance careers. And yet, by avoiding the task of identifying my own preferences and strengths and focusing only on my weaknesses, I was perpetuating the same narrow view.

One: *Those weaknesses being?*

Two: One is my so-called ‘foreign’ expression of music; another relates to my technical insecurities. As my work in the next chapter will seek to disseminate, my lack of both awareness and understanding of cultural and racial significance within Western classical music contributed to an inner tension that, I believe, stunted the growth and development of my musical voice and my identity. My technical insecurities, in particular, arose from formative experiences in competitions during my adolescent years, when I noticed a lack of rigorous technical training in my musical education. However, it was years later that I took the initiative to seek a teacher who demanded technical excellence as a prerequisite for musical development. My work with this teacher, who had many technically brilliant students (several trained in East Asia), introduced me to strengthening and training exercises, studies, and a deeper physical awareness of piano-playing. Yet certain comments, made during one of our first lessons together, made a

lasting impression. I was told that my musical level was very high, but my technical level did not match up. Such a statement, which confirmed what I already felt, was followed by the observation that Asian, and particularly Chinese, students were arriving at conservatoires in the West with highly advanced technical abilities. On the other hand, although Western-trained students usually could not compare technically, they were more developed musically. These comments added to my inner turmoil: I felt, at the same time, a foreigner in the West and lacking in the technical training of the East; I identified with neither.

One: *Okay, more of your work on East Asians in Western classical music will be discussed in the next chapter... come back to your transformation.*

Two: Yes, through my research and further musical development, I have come to terms with my appreciation for and desire to choose chamber music as my primary artistic focus. At the risk of seeming self-indulgent, I believe that this decision is a crucial step toward developing my identity.

One: *You say 'identity', but how does this fit in with your work on the cultural and racial issues? These represent parts of yourself that, as you mentioned, were a source of great tension.*

Two: I believe this relates to the public image of successful Asian musicians on the international performing platform; more often than not, these individuals have won prestigious international solo competitions and/or are touring soloists. Of those who

have made a career out of collaborative work (and I now focus primarily on pianists), figures such as Gerald Moore, Emmanuel Ax, and even Richard Goode come to mind, but rarely one of Asian descent.³⁸ On a more personal level, beginning from my time as an undergraduate piano student to present day, my interest in, and time spent on learning, chamber music has always been a curiosity to fellow East Asian pianists who felt that those precious hours could be better spent practising solo repertoire. This, combined with the view of those trained in Asia that chamber music does not have the same level of prestige as it does in the West, has shaped my idea of ensemble-work as the lesser option.

One: *So, you no longer hold such a view?*

Two: From the work that I've done and the experiences that I've had, I think it is no longer possible to hold such a view.

4.4 Speech in Practice: Germany, Autumn, 2018

The travel from London to southern Germany was not the most pleasant, nor scenic trip, but that was probably my own fault. I had been very stressed, as I usually am when it comes to transport by train. For some reason, I always seem to get on the wrong one, and this time was no exception. The journey to Berlin was fine, since I flew; I even stayed in the city for a few days with a close friend and caught up on some rest and socialising. However, the train journeys that brought me from Berlin to the chamber music festival near Frankfurt was

³⁸ The pianist, Wu Han, is one of the few that does come to mind and I've been fortunate to cross paths with her in the past.

fraught with the danger of getting on to the wrong train and being taken far, far away from my destination. This had, of course, happened in the past, when I was travelling within England to a competition; when all I had to do was ride the train for 10 minutes but, instead, took a service that nearly took me to Glasgow... This time, fortunately, I had only taken the right train in the wrong direction, and eventually managed to arrive at the festival. It was a warm, sunny day, and as I walked from the station to the campus, I was unusually aware of how loud my suitcase sounded, rolling over the cobblestones of the peaceful German village.

The repertoire that I had brought was the same as for France: Mendelssohn and Shostakovich trios. Musically, I felt adequately prepared for the ten days of rehearsals, lessons, and performances. The festival itself, however, was previously unknown to me and I didn't know what to expect; I had never worked with any of the faculty in the past, and aside from the friends in my chamber group, I did not know anyone among the participants. Thus, it was a fortunate and pleasant surprise when certain elements of my research began to pop up throughout the course. One particularly memorable instance occurred during a lesson with a violin professor; we had just finished playing through the first movement of the Mendelssohn, when the teacher began to discuss bowings with the strings. I rather liked these moments during 'coachings', when I could sit back and absorb information without having the pressure to immediately resolve an issue. At one point, the teacher stated that the expression of the violin melody, beginning at bar 22, could be informed by the sound of German speech. Immediately, my interest was pricked; I had never imagined that I would hear of speech prosody at this course, let alone at any classical music festival. The teacher began to impose a German sentence upon the melody and explained that the single-note anacrusis ('upbeat') of the phrase could be thought of as a definite article in German grammar (such as 'die' or 'das'). The significance of this was that the attack and length of

the note would be transformed; by thinking of 'die', for instance, the beginning of the note would acquire an attack associated with the voiced plosive consonant sound, 'd'.³⁹

Alternatively, a note could begin with a nasal consonant, such as 'm', or 'n'; or a vowel such as 'a'. As for duration, association with a grammatical article would require a separation between the anacrusis and the downbeat, preventing an elision between the first two notes of the violin phrase. In the case of this piece, the two notes are the same, so attention to the bow is required to make sure that they do not elide. As the lesson went on, the teacher continued to associate musical phrases with German speech, using his knowledge and fluency of the language,⁴⁰ and demonstrating on the violin when necessary. Timing suggestions were explained in terms of breathing and pauses, being rather like the observation of commas or full stops when reading a sentence out loud. Stresses on certain notes were explained by speaking German words in order to observe word-stress and then mimicked on the instrument. I knew that these directions were entirely applicable to the piano, and was ecstatic. I had formed similar ideas in my own practising of German music (see my work on the Berg sonata in Chapter One), and although the teacher's explanations were in terms of string-playing, including bowings, fingerings, and strokes, I could interpret the effects from a pianist's perspective, with wrist motion, finger touch, arm gestures, and pedalling.

This and subsequent lessons were hugely significant for me; I gained confidence, both in my research and performance, and felt justified in my obsession over speech and its connection to musical expression. I felt elated that a musician of such esteem was using

³⁹ These linguistic explanations are my own, and not those of the teacher.

⁴⁰ Interestingly, the violinist's native language is Catalan.

techniques similar to ones that I had discovered on my own; no longer did I feel that my ideas were foolish attempts at compensating for my expressive inadequacies. Finally, I had found validation in the musical practice of someone other than myself; little did I know that this was just the beginning.

Halfway through the course, a lecture was given by another member of the faculty; a violinist. The focus of the lecture was on Mozart's String Quartet no. 15 in D minor and the possible expressive connections between the chamber work and his opera, Don Giovanni. The talk was held in a small, carpeted room; too small, we all realised, as we slowly filled up the space. At the front hung a large screen which took up the entirety of the wall. The lecturer sat towards one side, fiddling with a laptop that was connected by a long cord to the screen. As I looked around the room, I noticed that no one seemed to have a concrete idea of what was going to be discussed; I, myself, was expecting a lecture on certain historical significances and/or musical elements that linked the two works. One of the pre-formed string quartet groups also sat at the front of the room, and from the look of the set-up, seemed ready perform the chamber work. As the audience sat and listened, excerpts from both the string quartet and the opera were introduced and performed (in the case of the opera, a recording was played). During one of the musical examples (the duet 'La ci darem la mano'), the lecturer began to compare the melodic phrasing of the vocal line in the opera with the opening phrase of the string quartet, which involves a main voice, played by the first violin, with textural accompaniment and bass provided by the three other players. At one point, a striking concept was introduced to us: a potential strategy for discovering how to phrase the first violin line in the chamber work is to observe how the Italian words of the text in the aria are sung. According to the lecturer, the text places certain restrictions on how notes are to be phrased; certain notes are connected to one another while others are

separated, according to the words that are set underneath. One example is at the beginning of 'La ci darem la mano', where Don Giovanni sings the words 'vedi, non è lontano': in the score, there are no slurs in the vocal line to indicate which notes should be grouped together and which should be separated. According to the text of the libretto, however, the first two notes (a quaver on 've-' and a semi-quaver on '-di') should be grouped together, as they form a single word ('vedi'), while the third note (a semi-quaver on 'non') is slightly separated from the previous two due to the comma before 'non' in the text. Obviously, these details provide a wealth of information for instrumentalists who seek to play a melody in a 'sung' or 'spoken' manner (as is often the goal, even where there are no words). Awareness of such information also affects the physical movements and gestures of instrumentalists, and enhances the communication and conviction of musical phrasing during performance. The lecturer also described the importance of rhyme in both text and music; by acknowledging that certain notes within one phrase are meant to rhyme with those in another, the performer can evoke the sense of rhyme through his/her musical expression. An example of this was 'la mano' in the opening phrase and 'lontano' in the subsequent phrase. Towards the end of the lecture, and in a further attempt to emphasise the significance of the sound of Italian speech, the lecturer engaged an Italian student to read an excerpt from an 18th century Italian poem to the audience in an effort to familiarise us, as the listeners, to the prosody of Italian speech. As the student spoke, I was struck by the rhythmic nuances of the words and phrases, and how distinct the speech sounded from English. I noticed the tendency of words to have CV-type syllable structures, with a distribution of long vowels in open syllables and short vowels in closed syllables. At that moment, the concept of singing as important to the performance of instrumental music took on further specificity for me: in

addition to the beauty of a singing tone and lyrical phrasing, the particular characteristics of the language being sung holds significant relevance.

I was eager to speak to the presenter about my research after the lecture; the connection between what was presented and my own work seemed undeniable. As I stood in the small line that formed of those wanting to commentate or discuss in further detail, I felt nervous. So much so that I almost left the line with the justification that I could speak to the lecturer at a calmer and more isolated moment. The thought of expressing my considerations and questions to a teacher whom I didn't know, and in front of a group that was not made of doctoral trainees and/or academic researchers, was nerve-wracking. Would I sound too eager? Would my words jumble up and become senseless ramblings? Essentially, I felt insecure about discussing my work with other performers; the personal struggles that led to my research interest seemed too revealing of my musical inadequacies, and I was afraid to have them known among the performance community. Eventually, it was my turn to approach and I briefly introduced myself and my research. Fortunately, the lecturer was interested, and suggested that we meet for lunch to discuss in further detail.

At lunch time, the lecturer, admitting to being more familiar with theoretical and/or impressionistic literature on the topic of language and music, expressed an interest in knowing more about current empirical studies. I responded with the bits and pieces that I could recall from my literature review, as well as gave an overview of my own experimental design. I introduced the work of Aniruddh Patel, with which the lecturer was unfamiliar, and explained how the contributions of Patel and others have informed my research journey. At one point, I began to feel slightly uncomfortable with the one-sided nature of the conversation; I seemed to be simply reciting facts, rather than engaging in discussion and gaining insight. Seeking a shift in dialogue, I began to speak about certain musical

experiences and concerns that led me to pursue such research. With this change, the lecturer began to contribute more as a teacher and as an observer of the changing cultural climate within Western classical music. I used this opportunity to pose concerns about authority in relation to language background. As I had expected, the work of Béla Bartók came up, with regards to the composer's own connection between the Hungarian language and Hungarian folk music (see the 'Harvard Lectures' in Suchoff, 1976; also Kidd, 2008). Other composers who expressly associated music with language include Zoltán Kodály (see the quote in Lesznai, 1961: 56) and Leoš Janáček (see Pearl, 2006). These names often appear in my conversations with classical performers when the topic of language is brought up; this is likely because the act of being aware of the background and influences of a piece of music is part of a performer's learning process. However, in my experience, performers' awareness of the relationship between language and music is often exclusively associated with the above-mentioned composers, or else simply mentioned in passing. Hoping to go further, I explicitly inquired about the lecturer's personal experiences; whether they knew of anyone who tended towards expressing music a certain way, based on language background. The response I received was that this was a very complex topic, involving numerous factors, but that there were a few observations that could be relevant. Once, the lecturer had attended a live performance of John P. Souza's popular piece, 'The Stars and Stripes Forever', played by a well-known and esteemed European orchestra. The lecturer's view was that, although the orchestra was entirely professional in its level and experience, the performance 'couldn't have been more wrong'. The reason for this feeling couldn't be explained; it simply generated an instinctive feeling of 'wrongness'. Sometime after this conversation, I read a bit of background on Souza's piece and discovered that the composer had written verses to the melodies! It may have been possible that the orchestra performed the piece without an

awareness of the English text (or even the significance of the English language), and the result could have been as 'wrong' as if one were to play the melody of 'La ci darem la mano' on an instrument, without considering the Italian text. The lecturer also observed that there may be certain instinctive qualities of expression within certain people that are based on their language experience, and that perhaps these qualities allowed for those to master certain expressive elements of music faster and with more ease than others. I agreed, adding that I believed in the possibility that those who were not born with 'certain instinctive qualities' could acquire them through exposure and/or education. The lecturer nodded in agreement but didn't add anything further.

At this point, I steered the conversation towards a more personal and sensitive avenue, explaining how my own struggles as an East Asian classical musician had been the predominant force behind my research journey. On this topic, the lecturer contributed the experience of an incident that was 'confounding to this day': during the break in a day of entrance auditions for a prestigious European conservatoire, a world-renowned musician on the panel exclaimed, exasperatedly, 'If I hear one more Korean violinist, I'm going to scream!'. This statement was disturbing for my lecturer, as it implied a particularly negative connotation toward one specific group of musicians. Moreover, one would never hear someone complain, 'If I hear one more Jewish violinist', or, '...violinist of colour', and yet it was somehow acceptable to say such a phrase about East Asian performers. My response was that statements like this showcased the necessity of raising awareness and discussion around the existence of social prejudice of East Asians within Western classical music. Upon further reflection, I also recognised the importance of voicing one's concerns as a researcher, performer, and/or teacher, as the likelihood that incidents like this are lost or buried under

societal, institutional, and/or professional pressures may be significant, particularly among students and young professionals.

I felt fortunate enough to have had a conversation on my research, but as luck would have it my chamber group was also offered a coaching session with this lecturer. We decided to prepare the second movement of Mendelssohn's C minor trio, Andante espressivo. During rehearsals of this movement, we had difficulty achieving a sense of freedom in our playing; we discussed that there seemed to be something missing, something that would allow for more openness in our expression, while still maintaining unity as an ensemble. Particularly, we struggled with the first section of the movement, which contains the first theme, in two parts. In each part, the piano plays alone and then the two strings repeat the material as a duet while the piano provides a gentle, rhythmic accompaniment. The theme itself seems undeniably vocal: the four-part voice writing in the solo piano sections creates a chorale-like texture, with the main melody in the soprano and the other voices weaving around, either in support of, or creating tension with, the highest voice. When the strings enter, the vocal quality of the melody is enhanced even more by long sustains of the bows and a voice-like quality of vibrato. Our awareness of these elements was helpful, but not enough to achieve the level of expressivity that we desired in our performance.

The lecturer sensed our discomfort and dissatisfaction as we played through the movement. When we finished, it was suggested that we record the session so that afterwards, we can hear for ourselves the transformation of our music-making as the lesson progressed. This, the lecturer believed, was much more useful than 'blindly' trusting the words and judgments of the teacher. With that statement, I felt that this lesson was going to be different from the traditional format; at least, there seemed to be an alternative dynamic between student and teacher. Perhaps this was due to the fact that the lecturer was a

violinist, and as such, had never played the repertoire we were presenting, and possibly even the instruments upon which we've spent countless of hours learning and perfecting. Of course, the viola, being a stringed instrument, has similarities to the violin and cello, but I still sensed an openness from the lecturer to seek sources of knowledge and inspiration outside of one's own instrument.

A significant source of knowledge, in the violinist's opinion, comes from observing 'great singers', particularly of how they breathe. When playing a melody, it is crucial to know where are the intakes of breath, and how long are the lines before another breath is needed. According to the violinist, the moment that a phrase has breathing in the places that instinctively make most sense, then the music 'has life' and is convincing to the listener. The concept of 'freedom' was another topic of interest during the lesson. As we worked through the movement, we discussed the presence of freedom in musical performance; how highly experienced and knowledgeable performers often achieve a sense of freedom, of openness within themselves, which then transfers to the audience. I considered instances when I felt that certain performers seemed to just do whatever they wanted, and the music would always sound 'right'. I realised that, often, I would limit my expressive decisions during performance for fear of making a wrong one, and then having to 'go back and fix it'. I felt safe in my attempt to limit my expression, because doing nothing seemed better than doing something 'wrong' and being thought of as 'foreign' or 'different'. At the same time, I envied those who seemed so confident in their expression, never having to worry about being out of place; even if they did something wrong, they gave the impression that it was simple to fix. I wondered why I couldn't feel the same. I also felt that there must be an answer to my problem; some solution, or formula, or way of understanding music so that I can go beyond my fears. I believed that knowledge was the key, and the violinist's words served to reinforce

my instinctive belief. 'First and foremost,' the violist said, 'freedom comes from understanding and realising the implications within the score; in other words, the structure [which, for the violist, is a loaded concept, one that involves a multitude of elements] and its implications for phrasing is the guide for one's expression'. Once the performer understands the 'map' of the music, which shows moments of tension, relaxation, breathing, and direction (either forward-leaning or slowly backward-pulling), then she or he can have the freedom and the confidence to commit fully to her or his expressive decisions during performance. The issue with our performance was that the listener could sense constraint in our expression; perhaps we did not know, with full confidence, the map of the movement, or perhaps we did not have the courage to communicate our awareness. Regardless of the reason, the result of such constraint was that the listener (in this case, the violist) became bored, because they were not being 'told' the music in a captivating way.

At this point, the violist introduced us to some 'tricks' of practice, gleaned from previous experiences and other performers, designed to aid in unlocking our expressivity. One particularly useful technique for the piano involved the rolling, or spreading, of chords wherever they'd occur within the music.⁴¹ The proposed theory was that, by rolling the chords that supported the melody, I could avoid a sense of verticality that is inherent in playing several notes at the same time. Also, by observing how the body feels in the moment of rolling the chords, i.e. how the hands, arms, and upper body move to achieve the desired roll, I was 'teaching' my body about the space required to create each harmony. Furthermore, when certain chords required more emphasis than others, I was aided by the

⁴¹ I believe the lecturer intended the rolled chords to be done exclusively in practice, although there is historical evidence for doing so in the performance of certain music.

manifestation of a physical and durational sense of emphasis, since rolled chords take more time to execute. This allowed me to rely less on a fast attack and louder dynamic for emphasis, and more on a physical, embodied sense of broadening, as well as a lengthening of time. In turn, this physical awareness of harmony allows for the melody, which becomes influenced by the chords, to have more free and variable timing, in 'the right places'.⁴²

A final, notable remark was on the subject of achieving a sense of unity in the ensemble. It was proposed that the sense of 'togetherness' came from being aware of and allowing for the different voices of a piece to respond to one another, rather than focusing on simply 'being together'. For example, rather than relying on the idea of a 'main voice' (which is often the highest voice) and 'supporting voices', it is important to acknowledge the symbiotic relationship between all voices: the main voice not only 'sings as a soloist', but must be constantly aware of the direction provided by the bass line; the middle voices do not exclusively follow (either the top or the lowest voice), but can either act as further impulse toward the direction of the phrase, or can counteract by pulling against the direction, creating dissonance and/or turbulence in the music. Thus, what happens within the ensemble of voices is determined by what can be understood of the score and of the structure of the music.

⁴² Expressive timing is not a matter of giving and taking time at random, but is governed by tension and relaxation, or stability and instability created by the harmony, among other factors. More on this below.

Dialogue 3: Agency and Authority as a Performer

One: The suggestion to record your lessons with ‘the lecturer’ is interesting. Immediately, this speaks to the concepts of agency and authority within Western classical music, but specifically relating to pedagogy and the student-teacher dynamic. Have you given more thought as to how the action of or even the allowance for recording lessons could impact one’s musical development?

Two: Agency and authority are highly complex social concepts and, unfortunately, I am unable to provide a thorough analysis on either here.⁴³ What I can do, however, is discuss some personal observations that I believe touch upon these concepts and, hopefully, promote further discussion. Authority and agency have been crucial interacting factors in my education. I worked under the one-on-one, private teaching model from the beginning of my musical training until my current situation within a higher-education conservatoire. This model supports the view that the teacher is the ultimate authority, and stems from a long Western tradition of musical training, firmly established since the Middle Ages (for further detail, see Davidson and Jordan, 2007). This view seems most effective at the beginning stage, when the student is developing the skills and knowledge necessary to play music on her or his instrument. As the student progresses, however, the authority of the *music*, rather than the teacher, may develop; it is at this stage that the student should be equipped with the necessary skills to

⁴³ For further reading on agency and its various theoretical debates, see Emirbayer and Mische, 1998.

understand music for her- or himself and thus develops the agency to seek further information, interpret the score, gain performing experience, etc. As this happens, the locus of authority may not necessarily shift from teacher to the score, but become broadened to encompass both. A complicating situation occurs, however, when other external figures within the student's life gain authority; in my situation, this was both my own mother as well as professional recordings. I won't get into the details of the first now, as I believe it ties in with my work in the next chapter, on 'East Asian parents in the world of classical music'. The conferment of authority upon the recordings of others, however, occurred at an early stage in my development and, I believe, was due to a lack of musical knowledge (e.g. harmony and structure), which made the progression to placing authority directly with the score very difficult. In my subsequent experience, this played a detrimental role in what I believe ought to be the last stage of a student's education, when she or he begins to embark upon the 'real world' and sustain a professional career in music. At this late stage, both authority and agency are ideally obtained by the performer, leading to the development of musical identity and voice. Of course, musical development rarely occurs according to a straightforward trajectory. Along the way, tensions between sources of authority, or some kind of imbalance, may arise: in my case, the number of external sources, combined with my inadequate musical knowledge and a lack of cultural 'belonging', resulted in the underdevelopment of my musical authority and identity, as well as limited my expressive agency.

One: *And how does all of this relate to the recording of lessons?*

Two: If a teacher suggests that a student records the lessons, or, as another example, sit in on other students' lessons, this potentially may dissipate the teacher's authority, or at least provide more agency for the student to listen, judge, and make decisions about their own musical practice. I believe that this may help to rebalance the distribution of authority, serving as a useful tool particularly for those struggling to cross into the later stage. At the very least, it stimulates reflexivity.

***One:** I suppose this also ties in with your observation about the varying preferences of different teachers during the masterclasses in Belgium.*

Two: Yes, that is also another reason why agency is important, particularly for advanced students; when faced with multiple opinions, some even contradicting, one must be able to make informed musical decisions for oneself.

***One:** Focusing now on the concept of musical expression in performance, there seems to be some inconsistency throughout the reflections regarding your own perception and understanding of it. Towards the beginning of your recollections, you described your revision process of the op. 142 Impromptus of Schubert and a moment of introspection in which you felt that 'the music' existed beyond any physical or bodily 'realm'. This notion was highlighted by your inclusion of the Italian cellist's story, as well as a past statement from your teacher. As the recollections progressed, however, other elements seemed to influence and alter your experience of musical expression, particularly in private practice. The element of speech and its potential to inform expressive performance decisions involving subtle variations of timing, loudness, colour, and more, was a significant*

discovery. You also mentioned, albeit briefly, the idea of the score as a map that indicated where certain expressive devices may be used to create emphasis, to breathe, to clarify grouping, to show tension and release, etc. Finally, during your lesson with the lecturer, you were shown certain ‘tricks’ that allowed for such expressive devices to be more convincingly produced; these tricks seem related to physicality and spatiality, of movement and the body, and encourage an embodied awareness of expressive gestures. Can you now elucidate your understanding of musical expression? Is it still the rather ineffable, ethereal phenomenon that was described at the beginning of your recollections?

Two: This seems the right question to ask, but the simple answer is that I cannot say that I have fully comprehended all the complexity involved in musical expression. However, I personally feel that I have travelled down a promising path, one away from the rather vague and unhelpful idea of music as universal and magical. As a practitioner of Western classical music, as a performer interested in music as communication both to and from others, I have felt frustrated by the mysticism surrounding the concept.⁴⁴ I believe that the complexity of musical expression should not preclude further awareness, clarification, and the gathering of practical knowledge. In my previous chapter, I made an attempt at clarification by theorising upon a multi-componential model of musical expression that allows for the various elements (inclusive, rather than exhaustive) to be displayed and organised for the sake of clarity. Rather than creating a definitive

⁴⁴ Refer to the Italian cellist’s story, described in Section 4.3, in which music is given universal and mystical significance.

representation of musical expression, I sought a simplified approach to dealing with such a complex phenomenon. As a performer, however, I am aware that, while beneficial, this model is not enough to translate knowledge to practice; more specific, practical tools are needed. The ones you so helpfully summarised in your question have been useful tools in my own musical practice.

One: *At your mention of performers, I'd like to bring up this statement that you made:*

'Essentially, I felt insecure about discussing my work with other performers; the personal struggles that led to my research interest seemed too revealing of my musical inadequacies, and I was afraid to have them known among the performance community.'
This deserves further comment; exactly what fears do you have that would prevent you from discussing your work with other performers? Do performers not form a significant portion of your readership?

Two: To address this, I will first refer to the reflections of conductor and autoethnographer Brydie-Leigh Bartleet (2009; Bartleet and Ellis, 2009) on this very topic. In her work, Bartleet discusses her feelings of vulnerability, doubt, and frustration relating both to her musical practice and profession as a conductor and to her autoethnographic writing. She describes fearing that her research was 'dragging her down', having a negative impact on her conducting and her relationships with her students (Bartleet, 2009: 11). This fear relates to the common concern among performing musicians of 'overthinking'. Bartleet goes on to quote Janet Ritterman from the latter's public lecture (2004): 'Most performers know only too well that thinking about what one is doing – or consciously monitoring what one is doing – at the wrong

time can easily be counterproductive'. This fear, that overthinking will be detrimental to one's music-making and performing career, may lead performers to shun the very idea of embarking upon a research degree. I myself have been conscious of this, although my fear has been less on how my research could negatively impact my own musical practice and more on how other performers might react to my research.

One: *But if your music-making has benefitted from your work, as you have implied, then why does it matter if other performers do not value research in the same way?*

Two: I suppose another reason for fear comes from the exposing of my vulnerabilities and struggles in my musical practice. As Bartleet and Ellis (2009: 10) recognised, the 'potential darkness of vulnerability that comes from revealing stories, lives, creative decisions' may be disconcerting for many musicians, as they are 'accustomed to keeping such personal characteristics and problems hidden from public view'. My own concerns relate to the potential judgements of others and even the possible harm to my own reputation as a performer. As witnessed in some of the interview responses discussed in Chapter Five, recognition and expression of one's own musical difficulties may result in doubt about one's talent and potential for a performing career (Yoshihara, 2007: 199–200).

One: *So then why does this thesis exist?*

Two: I have ultimately decided that, although there may be performers who doubt the relevance and usefulness of critical and personal reflection upon my own practice, there

may still be those who would benefit greatly from my work. As Bartleet's inner voice had to remind her in moments of doubt, 'this is useful for musicians who don't mind making themselves vulnerable and who want to grow' (Bartleet, 2009: 11). I have been one of those musicians, and I deeply concur.

Chapter Five: East Asian Musicians in Western Classical Music

5.1 Introduction

This chapter has been difficult to formulate for several reasons: i) prior to my research journey, I had never considered the topic of race to be significant to my life as a classical musician; ii) I had never looked—or rather, never given myself the opportunity to look—beyond the obligations of my practice and performance to observe the cultural and social discourses on East Asian musicians within the Western classical music community, of which I am of course a member; lastly, iii) through work on this chapter, I was made to reflect deeply upon my experience as a child of Chinese immigrants, and thus confront some highly sensitive topics which include my diagnosis of a serious illness and my mother's controlling behaviour towards my musical training (issues that, as will become clearer below, are interrelated). What follows is a narrative account of my research into the literature on East Asian (henceforth, in line with common American usage, sometimes simply 'Asian') classical musicians, a backdrop against which I reflect upon my own experiences. My goal at the start of this research was to gain more perspective on the cultural situation of Asian musicians, so as to locate my own identity within a particular group. What I've learned is that there are numerous situations, as varied as the experiences of each musician and each individual; that I do not fall into any specific category or set of experiences; that, furthermore, an individual may have multiple and/or conflicting identities and that one's cultural identity 'represents a symbiosis or a compound of several cultural allegiances' (Petkova, 2005: 52). Nevertheless, certain experiences resonated strongly with me and helped to reveal elements of my own complex cultural and social identity that had previously been hidden.

The following reflections revolve around literature and interview transcriptions provided primarily by two sources: *Musicians from a Different Shore: Asians and Asian Americans in Classical Music* (Yoshihara, 2007) and *Soundtracks of Asian America: Navigating Race through Musical Performance* (Wang, 2015). Due to constraints on time, and concerns about an already very mixed methodology, I chose not to conduct interviews with East Asian musicians as part of my own research, and have relied significantly on these two sources. Both Yoshihara and Wang were trained in classical music (as a pianist and violinist respectively) before pursuing academic careers in ethnic and cultural studies, and I found their work especially relevant: they too benefited from ‘insider’ perspectives. Motivated by their work, I have made my own observations throughout the following sections very consciously as both a practitioner-researcher, having conducted an empirical experiment with Chinese, British, and American pianists, and as a Chinese-American pianist myself. I divide the chapter into sections, each confronting a salient element of the common discourse on Asian performers of Western classical music. The sections appear in the order in which I discovered elements of self-identification while conducting my review of the literature. Due to the complexity of individual identity politics and the influence of multiple (and sometimes contradictory) discourses upon it, this chapter, too, instantiates a fluid and complicated process of reflection rather than rehearsing a process leading to a concrete conclusion.

5.2 A Universal Language?

As mentioned in the Prologue, one of my earliest struggles as a pianist was questioning my ability to ‘speak’ the language of the music I was playing. I don’t remember exactly when or how, but the concept that classical music is a universal language was lodged within me at an

early age. I had almost a blind faith in that concept, even as doubts about my own ability to communicate in the musical language began to form, born out of negative experiences at competitions, judgements from established musicians, struggles during practice, and regular harsh—albeit well-meaning—criticism from my own mother. I had the idea that, since music was a language for everyone, and I was struggling to speak it, then I must have been the problem. As time went on, I developed low self-confidence in my abilities as a musician, as well as further confusion as I learned that there are vastly different styles and periods of classical music that require different kinds of ‘communication’.

Trying to be fairer to my younger self, I want to point out that the discourse of universality in Western classical music that easily pervades the hopeful minds of younger music students refers mainly to the ‘unifying qualities, shared humanity, and unique capacity to transcend racial and national borders’ (Wang, 2015: 11). It is meant to encourage those, particularly in certain racial minorities, that they can be included and can benefit from the study and the beauty of classical music. Yet, even as the transcendent view of classical music remains the predominant discourse in society and in education, the historically based—and equally entrenched—understanding that Western classical music is the ‘embodiment of European essence, achievement, and tradition’ (Wang, 2015: 12) still looms large, if in the shadows. It is this shadowy view that presented itself more to me than the romanticised and optimistic version. I came to understand myself as an outsider, one that sought vainly to play Western classical music with the same innate understanding as Western musicians did.

As Yoshihara (2007) clarifies, East Asian musicians’ own understanding of the relationship between cultural/racial identity and music-making may be categorised as two views: the universalist and the particularist. From numerous interviews conducted with

Asian musicians in the U.S., Yoshihara noted that those with universalist views believed that i) it is racist or at least provincial to think of musical understanding as geographically and culturally bound, ii) no one owns classical music any more legitimately than others, iii) in today's society, with easy and immediate access to information, it is possible to gain understanding of other musical cultures, and iv) the ability to express musical ideas, including those of cultures other than one's own, is a sign of a competence and talent (summarised in Leppänen, 2013). The particularist view, on the other hand, sees the nature of musical expression as culturally specific. It assumes that i) music-making is shaped by one's racial and cultural background, and ii) if the performer and the composer are of the same nationality or ethnicity, the music-making may be considered more authentic (more on authenticity to follow). Both perspectives have their obvious dangers at either poles: the universalist position can be charged with lacking determinant historical and social specificity, whereas the particularist position can be criticised for supporting racial essentialism (the idea that a culture has a set of characteristics that is exhibited by everyone in the community, while the diversity among different members of the community that have contributed to development of such characteristics are ignored; Yoshihara, 2007: 191). Words such as 'natural', 'innate', and 'instinctive' are liberally used within particularist discourses. According to Yoshihara, however, much of Asian musicians' experiences and ideas as performers show that they do not function within a dichotomy of either/or; rather, they are able to address the concept of authenticity in their own terms: by being faithful simultaneously to the music and to themselves (Yoshihara, 2007: 192).

I must mention at this point that, when I first read these words, I paused momentarily to consider whether I fell into the category of Asian musicians who seemed to have overcome the prevailing discourses to find their place in the world of Western classical

music. I decided that, no, I did not; as my previous chapter describes, I have been dealing with doubts and insecurities about my musical voice and expression for a significant proportion of my musical life. These doubts, shaped by certain experiences, as well as a lack of awareness—or at any rate acceptance—of the complexity of my cultural identity, has been the impetus for my research journey. Early on in my research, I seemed to follow the particularist viewpoint; my reasoning was that, as a performer, I placed value on the study of the unique ‘musical language’ and background (musical, cultural, and linguistic) of the composer. As my research progressed, however, I became more aware of how complex the interaction between composer, composition, performer (and even listener) within the experience of musical expression can be. My preference for acknowledging the differences between musicians from varying language backgrounds does not necessarily equate to an illiberal tumble into the pit of essentialism. Rather, and through both research and musical practice, I have gained a more nuanced and complex view of expression in Western classical music. This perspective allows for the coexistence of differences between individual musicians and differences between language-groups; this perspective also acknowledges that each individual may have their own set of experiences, knowledge, and abilities that affect their musical (and linguistic) expression to varying degrees. In other words, expressivity should be viewed as an individual, fluid and *malleable* ability, influenced but not delimited by language and culture, rather than a fixed capability.⁴⁵

⁴⁵ Musicality and expressivity are often said to be innate and impossible to teach. It seems to me that this could lead to instructors giving up on certain students, having concluded that they are not worth the effort; it is impossible for them to learn. Neuhaus (1993) writes about the act of teaching students who are ‘average’ compared to those with great talent and ‘genius’. He believes that, ‘in spite of the mysterious nature of the gift of “genius”, it is possible not only to describe it, but also to study and analyse it.’ He continues, ‘I believe that

But what does this mean for the formation of my own identity as an East Asian classical musician? I will address this further in a later section, where I will detail personal experiences and cite specific interviews with musicians whose responses resonated deeply with me. Before this, I will explain in further detail certain aspects of (musical) society that seem to contradict the universalist viewpoint.

5.3 Racial Marking and Stereotyping

Reading the texts of several sources on East Asian musicians in classical music, it is apparent that racial marking within the media has played a significant role in perpetuating a sense of 'otherness' with regard to East Asian classical musicians.⁴⁶ With frequent reference to the success of Asian musicians in the media as a sign of dominance and a 'taking over' of American or European orchestras and higher education institutions, there is an implied danger that Western classical music will be forcefully taken away from Western cultures by East Asian musicians. Leppänen's (2013) survey of discourses within three classical music

by attempting to the best of our ability to fathom the "mechanics" of the highest musical gift we shall always extract something useful that can be applied even to the most average pupil.' Finally, he mentions the behaviour of some of his colleagues: 'More than once my colleagues, hearing me teach, have hinted to me that I was being quixotic, that all the same it will never "come out" the way I want it.' While Neuhaus's words may seem blunt and outdated in terms of dividing students into 'genius' and 'average', nevertheless it is a reminder of the responsibility of teachers to pass down not only technical and theoretical knowledge, but to nurture artistry and expressivity.

⁴⁶ This concept of the 'Other' comes mainly from Edward Said's work on orientalism (Said, 1978), in which he criticises the Western view of the Middle East. By extension, it has been used by scholars to discuss the West's view of the Far East.

magazines in Europe from 2002-2011 determined several categories that frequently appeared in published articles, the first being the discourse of imminence, which was interpreted based on threatening, 'culturist', and/or warlike lexicons.⁴⁷ I myself remember being guilty of perpetuating this view of 'invasion' during my undergraduate years, when I'd childishly utter phrases like, 'Asians are taking over [the institution]'. My friendship circle at that time was made up of non-Asian musicians.

Additionally, Yoshihara views the constant reference of ethnic and racial categories in the media as neglectful of Asian musicians' individual origins, upbringings, cultural identity, and respective relationship to Asia, the United States, and Europe (Yoshihara, 2007: 5). This perpetrates the notion of 'sameness' with regards to Asian musicians, which not only hinders Western society's recognition of East Asians as having a range of national identities (let alone individual identities), but may even prevent the musicians themselves from recognising their own unique and individual circumstances. One interviewee in Wang (2015) mentions an increasing awareness of a 'subtle and yet steady undercurrent of resentment' at Juilliard, where he obtained his Master's degree:

There's a lot of negativity towards Asians [...] I felt like, okay, you know, there's another Korean kid, there's another Japanese kid, and then I felt like some of the professors felt that too and just assigned basically numbers to us. It wasn't anything necessarily personal. And by having so many Asians around, too, walking down the hallway, you're just another one, you know,

⁴⁷ Interestingly, out of the three magazines from Germany (*Crescendo*), United Kingdom (*Gramophone*), and Finland (*Rondo*), only *Gramophone* and *Rondo* featured articles on the so-called 'invasion' by East Asian classical musicians; the German magazine avoids mention of the phenomenon.

and there's not really a way to distinguish yourself from other people unless you're given the opportunity to perform (Wang, 2015: 78).

Furthermore, Wang (2015) notes that, while her interviewees were critical of the racial stereotypes of being robotic, technically proficient, but unemotional or insufficiently sensitive to the music that they play, they eventually come to internalise, accommodate, and even authenticate whatever partial truth there may be in such stereotypes. The interviewee quoted above observes that just as the 'majority of non-Asians believe' these stereotypes, he tended to believe them too (Wang, 2015: 79). Regarding the Asian musicians who seem to embody such negative clichés, he adds that, 'I think it's not a fault of their own; it's a fault of their training'. Exactly what this 'training' entails is beyond the focus of this thesis; however, I must admit that I had similarly vague notions of the difference between Eastern and Western instrumental training. My belief was that Eastern training focused predominantly on technique and discipline, rather than on musicality and artistry, which was more of a Western focus; I even felt in some ways musically superior, since I grew up and was trained in the West.⁴⁸ Looking back, I feel ashamed at my own contribution towards the perpetuation of stereotypes that were not based on any knowledge beyond what was casually repeated within my musical environment. Consequently, by acknowledging the partial truths of these negative stereotypes while simultaneously believing that I did not demonstrate such flaws, I functioned in a suspended realm between identifying myself as Asian and non-Asian.

⁴⁸ This feeling was further encouraged by numerous mentions throughout my musical development from teachers, fellow students, and audiences that I was 'actually an emotional and expressive performer'; that I was 'different from other Asians' and essentially 'didn't sound Asian' when I performed.

5.4 The Model Minority

Of course, not all stereotypes of East Asian musicians are negative, at least not on the surface. Asians' over-representation in institutions of higher education, not just in music but in other professional fields, have given them the title of the 'model minority'. Yoshihara (2007: 4) defines this designation as 'those who rise in the existing social structure through hard work and attain success in Western culture without posing a direct challenge to the economic and political status quo'. Successes of Asian classical musicians are often attributed to traits presumably specific to Asian culture, such as work ethic, family values, and the investment of time and money into education. As a result, Asian musicians' successes are reduced to a cultural essence that is shared, it is presumed, by the entire racial group. Yoshihara (2007) criticises these claims, stating that it is impossible to compare, or even quantify, values and traits among racial or ethnic groups. This is because one's cultural, educational, and economic success are conditioned by the availability of opportunities and resources in one's immediate environment.

I investigated this topic, discovering that the 'model minority' image may be a product of the interaction of several related elements. I began at the beginning, with a historical overview of the impact of Western classical music upon East Asia; this provided a foundation for understanding the mentality that many Asian parents have towards discipline and their children's success. From there, I considered both Western society's view of Asian families, as well as Asian parents' own views of the Western cultural framework and of classical music. As a child of immigrants, I was raised with certain pressures and experienced, first-hand, the tensions between my parents' views and the contrasting environment of school and society at large. Thinking back on my childhood in the United States, I now recognise many instances when my parents, and particularly my mother,

struggled to accept—or simply clashed head-on with—certain mind-sets and behaviours. As I will discuss later in this section, the difficulty of settling into a new country and culture, along with significant language barriers as well as financial insecurity, has resulted in many Asian parents (my own among them) reacting with a seemingly overzealous desire for the success of their children; furthermore, this success is often believed to be achieved only through ‘Eastern’ values of hard work, unswerving dedication, and sacrifice. To these parents, Western classical music training became more of an Eastern pursuit: a representation of Eastern, rather than Western values.

History of Western Classical Music in East Asia

To contextualise East Asia’s interest in Western classical music, as well as the discourses surrounding East Asian parents, it is necessary to understand how classical music was first introduced in East Asia and the reasons for its extraordinary impact. Most literature focuses on the three countries that have produced the largest of number of Asian musicians within the field of classical music: Japan, Korea, and China.

Western classical music first arrived in the three nations through three institutions: the military, churches, and schools. Japan first opened its doors to diplomacy and trade with Western nations in the 1850s, resulting in the modernisation of its military through the observation and study of Western armies. One observation was of the importance of military bands for boosting morale and maintaining discipline in the military. These bands (first founded in 1871) introduced elements of Western music, including orchestral instruments and piano, Western singing style, and the public concert (Wade, 2005: 12; also see Eppstein, 1994 quoted in Yoshihara, 2007: 16). Likewise, in China, leaders discovered the capacity of Western music as a tool for motivation, organisation, and discipline, and

reformed China's own political-military culture accordingly. After the Sino-Japanese War in 1895, aggressive reforms of the China's government and education included the creation of its first Western military band, with subsequent ensembles forming in various cities around the nation (Melvin and Cai, 2004: 84-86). Korea established its first military band in 1881, following observation of European, American, and Russian naval bands around the area of Jemulpo Harbour, where they were often based (Hwang, 2001: 75). These associations of Western music with modernity, discipline, and geopolitical power were significant influences during East Asia's introduction and early adoption of Western music (Yoshihara, 2007: 16).

Another influence of Western music in East Asia came from Christian missionaries and missionary schools. In the late 16th century, Jesuit missionaries presented the Chinese Emperor with a clavichord and gave lessons in Western music at the Imperial Palace in Beijing. Missionaries arrived in Japan in the mid-16th century, bringing with them Western music, although it wasn't until the mid-19th century that Japan fully opened its doors to Western influence and the Christian hymns and hymnals of Protestant missionaries were introduced (Mehl, 2013; also see Harich-Schneider, 1973). Protestant missionaries also arrived in China in the 19th century and introduced organ music and choral singing, exposing many, especially in rural areas, to European music (Melvin and Cai, 2004). In the late 19th century, missionaries introduced Protestant hymns to Korea (Hwang, 2001: 76-78), teaching Western concepts of part-singing and sight-reading in mission schools, and familiarising schoolchildren with the melodies, harmonies, and structures of European music.

Alongside mission schools, Western musical education was incorporated into secular school curricula through the 'school song' movement, active in all three countries; this allowed for Western music to be 'indigenised' across the nations, by way of the learning of

art songs and popular songs in primary schools (Yoshihara, 2007: 17). In Japan, songbooks for schoolchildren were compiled by combining Western tunes with original Japanese texts, composing new Japanese songs using the Western scale, or harmonising traditional Japanese songs using the Western tonal system (Wade, 2005: 12-14). China, following its defeat in the Sino-Japanese War, made use of Japanese songs in its own primary education system; Chinese intellectuals and students who travelled to Japan brought back these songs and set them to Chinese lyrics (Melvin and Cai, 2004: 86-88). In Korea, songs that evolved from Protestant hymns, or *ch'anggas*, circulated throughout elementary education; new art songs and popular songs modelled after *ch'anggas* emerged in the following decades (Lee, 1988: 12-13).

Beyond primary education, the three nations also established institutions of advanced musical training. Japan's Tokyo Music School and, subsequently, the Music Department of the Tokyo University of Fine Arts and Music, began as the Music Investigation Committee during the Meiji period (1868-1912); this committee was responsible for exploring ways to incorporate Western music into Japanese education (Eppstein, 1994: 28-30). In Korea, the first department of Western Music was established in 1910 at Ehwa Women's College (later Ehwa Women's University) in Seoul. In 1912, the Chosun Institute of Classical Music was founded as the first Western-style music institute. Subsequently, the School of Music of the Seoul National University was established (1946) and the Korean National University of the Arts became the first conservatory-style musical institution in Korea (1993; see Lee, 1988 and Hwang, 2001). In China, the first conservatory of music was built in 1927 in Shanghai; following the founding of the People's Republic of China in 1949, the Central Conservatory of Music was established in Beijing (Melvin and Cai, 2004).

As Yoshihara (2007: 18) articulates, Western music's influence moved not only from the West to the East, but also within East Asia by way of power relations and regional politics. Japan, during its period of military and economic expansion, influenced China and Korea to modernise their educational systems, while becoming an appealing destination for Chinese and Korean students to study abroad. Chinese musicians and intellectuals educated there (and in Europe and the United States) went back to China to teach and create a new form of 'national music' that combined Chinese melodies with Western harmony and musical forms (Melvin and Cai, 2004: 92-95). After 1910 and Japan's annexation of Korea, serious Korean music students went to study in Japan; Young-Whan Kim, as the first Korean musician to study abroad, returned to Korea in 1918 to teach at what is now Yonsei University (Lee, 1988).

Struggles between the three nations due to colonialism also influenced the impact and cultural meaning of Western music. In the 1930s, as Japan exercised increasingly aggressive colonial power in China and Korea, Western-style music became a form of resistance *against* the Japanese. After the bombing of Shanghai by Japan in 1932, a musical movement emerged in the city that involving the creation of unaccompanied songs to be sung by the masses in resistance to Japan and in sympathy with those poor and devastated (Jones, 2001). Music became revolutionary; more than just art or even a tool for social change, it was 'a weapon for liberating the masses' (Melvin and Cai, 2004: 126). Japan also saw a shift in the ideas and policies surrounding Western music at home; by the 1930s, the increasingly nationalistic demand that Japanese musicians perform Japanese music, along with demands for music that could reach the masses, led to the 'establishment of national music' and a 'cleansing' of Japan's musical culture of foreign influence (Yoshihara, 2007: 21, quoted in Akiyama, 1976: 409). American and British music, in particular, were considered

‘corrupt’, whereas the music of Germany and Italy (Japan’s allies) were ‘healthy’ (Yoshihara, 2007: 21; quoted in Akiyama, 1976: 554-56). Interestingly, Protestant hymns were filed in the former category.

At this point, it is necessary to discuss the particularly dramatic and tragic state of Western music in China under the Cultural Revolution, beginning in 1966. However, for the sake of fuller context, I will first mention the significant influence of foreign exiles upon development of Western music within China since the early 20th century, leading up to the Cultural Revolution. During Czarist Russia’s increasingly anti-Semitic rule, many Russian Jews fled to Manchuria; after the Russian Revolution in 1917, White Russians were likewise forced to flee their country. These very different exiles first moved to Harbin (then part of Manchuria) in north-eastern China, bringing their culture and their music with them. Prior to my research, I was unaware of the extent of influence that these Russian exiles exerted on cultural life in Harbin, the city in which I was born. My mother had, from time to time in my childhood, told me about the Russian population that once had flourished in Harbin, and their influence upon the food and architecture of the city. However, I had no distinct knowledge of the rich musical life that was established, including the First Harbin Music Academy, a symphony orchestra, and a string quartet (Melvin and Cai, 2004: 101). When the Japanese invaded Manchuria in 1931, these Russian exiles were forced to move to Shanghai, where they were joined later by European Jews fleeing the Nazis, beginning in 1938. By 1941, approximately 267 professional artists were living in Shanghai; several of these musicians were recruited by Xiao Youmei and his successors to teach at the Shanghai conservatory, established as the first in all of China in 1927 (Melvin and Cai, 2004). These exiles created the beginnings of a vibrant musical community that went on to flourish across China and have a significant influence on the country’s musical culture (Yoshihara, 2007:

20). Thus, it was a tragedy on a national scale when Mao Zedong, attempting to recover control of the Communist Party after the Great Leap Forward of 1958-62, launched the Cultural Revolution in 1966. This, along with destroying the customs, thinking, and culture of the 'exploiting classes', created a nation-wide persecution and ban on Western classical music and its performers (Melvin and Cai, 2004: 230). Instruments were confiscated or destroyed, scores were burned, and anyone who criticised or opposed the Communist Party was either killed or forced to abandon their profession and do manual labour. During this time, many conservatoire professors, spouses, and students broke down under mental and physical torture or committed suicide. Some of those fortunate enough to escape, including Central Conservatory President Ma Sicong, fled to Hong Kong and were eventually gained asylum in the United States (Calvin and Mai, 2004: 225-64).

After the death of Mao Zedong in 1976, accompanying the end of the Cultural Revolution, Western classical music was able to make a strong resurgence in China. Deng Xiaoping, Mao's successor, allowed conservatoires to reopen in 1977, and in 1979 Isaac Stern travelled to China, a visit that was documented in Murray Lerner's *From Mao to Mozart: Isaac Stern in China* (1979). This visit was one of the first high-profile efforts to re-establish cultural relations with the West after the Revolution. Twenty years later, when Stern returned to China, the documented visit showed a vast development of Western classical music within China; children showcased in the first film had become world-renowned musicians, and students at the Central Conservatory played more substantive works at a higher level (Miller, 2000; in Yang, 2014: 73).

I feel that I must offer some explanation at this point: I have focused more on China's history, compared to those of the other East Asian nations, because of my own personal background as a Chinese-born immigrant, raised in the United States. Growing up,

I never maintained a close connection with China; having left at the age of five, I never once returned to my birth city. Beyond the blurry flashes of early childhood memories and the speaking of Mandarin at home, I had little conception and understanding of China, and never had a chance (or, if I'm honest, felt the need), to develop further awareness. Yet I never could identify fully as an American; I still felt rooted to China in a way that is difficult to describe. I suppose that having some memories, no matter how vague, and knowing that I was born on different soil, kept a sliver of connection alive. I must admit, however, that I'd been ignorant of China's long and intricate history, as well as its complex relationship with the West. As for Western classical music, I never took the time to understand how such a vast and rich European tradition became the phenomenon that it is in the East. Reading about the musical culture that developed in Harbin, however, I felt for the first time in my life that classical music was, somehow, a part of my background and history; that the culture brought by exiles in the early 20th century managed to survive and be revitalised after the Cultural Revolution, and then was passed on to me. It is a silly and romantic notion, but also the closest feeling to belonging in classical music that I've ever had. And belonging, as this project has sought to demonstrate, is fundamental even to things that—at a personal or national level—are acquired.

Returning to a broader view, the mid-20th century saw Western music in East Asia shift from solely being in the hands of governments and intellectuals to spreading throughout the middle class; this was due to various sociocultural aspirations, by no means unique to the Chinese context, of course, that led to the transformation of classical music from a privilege of the elite to an art form of the middle class (Yoshihara, 2007: 33).⁴⁹ The

⁴⁹ For literature on this phenomenon in the West, during the 19th century, see Weber (1975; 2008)

growth in manufacturing of musical instruments for the East Asian domestic market also contributed to this change, as it spurred commercial interest in classical music as a growing industry. Additionally, the development of music pedagogy allowed for classical music training, particularly in the form of private instruction, to be more widely available. One instrument in particular became an appealingly solid sign of modernity and class aspiration: the piano. Initially an imported instrument that only elites could afford to own, national production of pianos began to increase dramatically in post-war Japan due to the technical innovations of two major manufacturers, Yamaha and Kawai. In 1954, Japan's piano output was around 10,000; by 1963, this number had risen to 290,000. By 1970, Japan had surpassed the United States and become the number one producer of pianos world-wide (Yoshihara, 2007: 33). Similar patterns occurred in Korea and China; by 1985, Korea became the world's third-largest producer of pianos after Japan and the United States. In China, piano production between 1991-1995 and 1996-2000 increased by 100 percent (Yoshihara, 2007: 34). Several factors seem to have contributed to the growing popularity and demand of the piano and of classical music in general: the introduction of instrumental training in music education, the growing popularity of music due to radio and television, manufacturer's efforts to increase demand, increased leisure time and recreation of the middle class, and the growth of personal income (Maema and Iwano, 2001: 245; in Yoshihara, 2007: 34). Increased advertisement of the piano as a symbol of sophistication and class also fed the growing demand, as well as the launching of music schools that provided private and group lessons for children at reasonable prices. Eventually, Western classical music became a form of cultural capital for the East Asian middle class. As the increase of access to musical instruments and instruction allowed more people from a range of socio-economic strata, including those with little to no background in Western culture, to

take part in Western classical music, it became a source of cultural advancement and upward mobility. Furthermore, the musicians trained within Asian systems began to attract world-wide attention, winning major international competitions, entering prestigious Western institutions such as Juilliard for further studies, and performing around the world. The association of Western classical music with modernity and the West, combined with the prestige and success of those who, through their music-making, developed international careers, made a deep impression on the world-view of the East Asian middle classes. As the next section seeks to detail, Western classical music as cultural capital and a claim to class identity has had a particularly strong effect on the world-view of those who strive towards success not only for themselves but for the next generation: East Asian parents.

East Asian Parents

Before detailing the work of Yoshihara (2007) and Wang (2015) on East Asian parents' roles and contributions toward the success of Asian musicians within Western classical music, as well as to the perception of 'model minority' within Western society, I want to provide some context by offering my own experiences as a child of Chinese immigrants living in the United States. The following section has been especially painful to write, but necessary: it provides the backdrop to my critique of certain Western discourses on East Asian musicians and the parents who support them.

My Story

I was introduced to the piano by my parents when I was four years old and still living in China. The few memories I have of learning the instrument include trekking over icy sheets of

ground during the wintertime in Harbin, holding onto my mother's hand, to arrive at a small candlelit room in which an upright piano had pride of place. I do not remember enjoying the lessons, as they seemed always to be held during the night and the yellow glow of the candles in the cramped space made me feel sleepy. However, I seemed to have made some impression upon the teacher, because he encouraged my mother to continue my training, which lasted for about a further year until we immigrated to the United States.

Our move to the U.S., and subsequent new life, was difficult for my parents in many ways; financially, we had very little, and it was a struggle just to decide what was the cheapest food to buy, never mind what toys I could have. My clothing came from Salvation Army donations, our house was owned by my father's employer and practically falling apart, and the only recreational activities we could afford to do was to fish in ponds or take long walks in the countryside. I did not have any more piano lessons until I was six and a half, after my father had found an old electric organ at a yard sale. From that point, my parents somehow gathered enough money to afford a piano teacher, upgrade to an old and tattered upright, and my life as a piano student was temporarily re-established.

I say 'temporarily' because two years later, at eight years old, I began to have severe pain in my legs at night. My father initially assumed they were growing pains, or perhaps a symptom of calcium deficiency, but the pain was severe enough that for several days, I would leave my bed in the middle of the night and lie in front of the TV to distract myself, all the while chugging down large amounts of milk. I also began to have severe nosebleeds, sometimes lasting for five minutes, and my colds would last much longer than usual. One night, my mother came downstairs to find me lying, pale and weak, and in so much pain that I couldn't even focus on the TV. Sensing that something was severely wrong, she insisted to my father that I was taken to hospital. I still have a memory of that experience, of the

moment before I lost consciousness in the emergency room—I had found it strange that I was made to wear a funny hospital gown—as well as the moment when I woke up. I was still in the funny gown, and was lying in a hospital bed, alone. Some of my clothes and possessions were stacked around the room, and I called out for my parents but no one came. I stepped down from the bed and walked out into the hall with bare feet. Finding a nurse, I was told that my parents were just in a conference room nearby, and I walked there but the door was closed. Eventually, my parents came out and brought me back to my room, saying that I was going to have to stay in the hospital—I had been diagnosed with Acute Lymphocytic Leukaemia.

A form of cancer, this disease required long-term hospitalisation as well as intense and physically harsh chemotherapy. Suddenly, the life that my parents were slowly struggling to establish in a new country was turned upside-down. It was years later when my mother told me of the only moment when she ever saw my father, a man who never seemed to show much outward emotion, cry so much—it was during their drive back to our house to bring more of my belongings to the hospital. The car was filled with my father’s wailing screams as he struggled to drive both himself and my mother home.

My mother quit her job following my diagnosis and lived with me in the hospital, caring for my needs and tending to the painful side-effects of my treatment. My father, becoming the sole provider for our family, was forced to work constantly: he barely had time to visit me in hospital. Fortunately, we were provided with low-income governmental aid for the costs of my treatment, and eventually—around two years later—my cancer went into complete remission. Around this time, I began to have piano lessons again. As my health recovered and the years went by, my musical practice and training increased in intensity.

At this point, it feels important to stress the fact that I had never chosen to learn the piano, unlike many of my fellow musicians in the West. I did not experience 'the spark' or eagerness to reach for the keyboard and spend hours delighting in the beauty of its sound, as others have described. I was, effectively, told to practise, taken to lessons, and made to improve my skills by my parents and eventually by my own feelings of guilt over their sacrifices of time, energy, and money. That I developed a deep and unwavering love for and devotion to classical music was, in my opinion, my good luck: I could just as easily have come to resent it bitterly.

My goal in providing the above context, with its highly personal and sensitive nature, is to relate my own personal background with those detailed in current scholarship in this area. Of course, each individual goes through experiences that are unique to their own situation and circumstances, and I did not expect, in my review of the literature, to find details mirroring my own. However, as I read the words and stories of other East Asian musicians, as well as of the parents who struggled to provide musical training for their children, I was able to gain a broader perspective, both on my personal situation and on the actions of my parents. Furthermore, through reflecting upon my own upbringing, I realised that the influence of my mother's caring discipline and controlling behaviour during my childhood and adolescence had a profound effect upon my personal and musical development. Specifically, her devotion both to my musical training and to nursing me through serious illness brought the issue of parental sacrifice into uncomfortably sharp relief (more on this to follow).

East Asian Parenting

It is widely known that East Asian families can be particularly driven when it comes to their offspring's education. Western society tends to view Asian and Asian immigrant parents as overbearing, excessively ambitious, and even draconian (see Chua, 2011); these opinions seem to stem partially from the reality (as far as my own experiences are concerned), but are also confounded by a lack of ethnographical understanding. Certainly, as a child raised in the United States, I viewed my own parents as having values that were extreme, and diametrically opposed to those of my classmates' parents, who encouraged and emphasised their children's right to have fun and make choices based on what they felt and wanted. I remember feeling frustrated at my lack of choice, at the discipline that prevented me from spending time with friends.⁵⁰ However, at other times I felt lucky that my parents were relatively inconsistent with their discipline; that, apparently, other Asian children had it much worse. Since I didn't have close friendships with Asians throughout my childhood and adolescence, my general understanding of Asian success in relation to parental discipline actually formed from a Western perspective: Asians are successful because their parents sacrifice, push, and demand excellence; end of discussion. Through my research, I became exposed to work that sheds light on the historical, cultural, and personal circumstances of East Asian parents, particularly of immigrants in the United States. Interviews with Asian mothers during Saturdays at the Juilliard Pre-college programme (see Wang, 2015) reveal the complexity of Asian parents' connection with Western classical music, as well as with a

⁵⁰ I acknowledge that music is not the only pursuit that limits a child's free time, as dance and sport activities may place even more limits upon a child's social experiences, as well as physical and mental demands. Furthermore, these activities are clearly not exclusive to Asian households.

new country, new language, and cultural framework significantly different from that of their own upbringing. Having access to the individual experiences and struggles of immigrant parents, spoken from their own perspective, influenced me to confront my own childhood and adolescent experiences and seek to understand the behaviour of my own parents.

First, it is necessary to discuss the notions that all Asian parents are competitive, protective, ambitious, and self-sacrificing. I find that the issue lies not in these characteristics themselves: having drive and commitment towards a child's education and success are obviously not inherently negative traits, nor are they exclusively Asian.⁵¹ However, problems arise when society emphasises certain behaviours of a group of people, while neglecting to understand the complexity of their situation and the reasons for their beliefs and behaviours, both as a group and as individuals.⁵² In the case of Asian parents, society's attention on their overbearing and controlling behaviour creates a generalised view and a superficial understanding of the East Asian experience through established tropes. Discourses on Asian classical musicians that emphasise the 'Asian' (or 'Confucian') values of discipline, obedience, and filial piety (see Day, 1994; also see Roth, 1997) have deemed it appropriate to generalise, and thus ignore the highly personal and nuanced filial experiences of Asian musicians. These discourses are further problematic for their implications towards the generalisation of Asian musicians as disciplined to the degree of being robotic and lacking in both emotional maturity and creativity (see Morley and

⁵¹ See the controversial work of Kim (2016), in which she details the similarities and differences between Jewish and Asian parenting.

⁵² See D. (2011) for a nuanced and personal examination of the 'model minority' myth and its harmful effects upon Asian-Americans.

Robbins, 1995). To supplement the lack of exploration into personal and nuanced experiences, I offer more of my own.

My mother was never musically trained, but she admired and valued the arts and always wanted to be a musician. During the early years of my own training, I viewed the fact that whenever I practised she would sit next to the piano and try to help as an extent of her love and passion for the life that she always wanted to lead. In retrospect (and I will discuss this further in the section below), I see this may have been her way of escaping from the difficulties of having to navigate a completely new environment, culture, and language since moving to the United States. As the assistance that accompanied me as a child (both in music and during my illness) transformed into a need to help me achieve the highest level of mastery as a pianist, following the decision that I would become a professional,⁵³ her comments were characterised by more obsession and urgency. During this time, I began to rebel. Through the inevitable development of my own self-awareness and need for independence, I realised that there was something wrong with my situation at home. Each day, I would go to school, where I focused on non-musical activities and spent time with American students who did not play any instrument and thus could not understand my experience, and after school I would come home and sit at the piano, only to get into heated arguments and fights with my mother over how I should practice, as well as over her judgements about my music-making, and my lack of discipline and/or obedience.⁵⁴ This pattern persisted, with more or less intensity, until I went to conservatoire in Boston at the

⁵³ This was made by my parents and piano teacher, and subsequently agreed by me when I was fourteen.

⁵⁴ My father would try to intervene periodically by claiming that I should be learning on my own, but to no avail, as my mother took that as a sign that he did not care enough about his own daughter.

age of eighteen. During my undergraduate studies, my mother even insisted that I go home during the weekends to practise with her, because she did not trust my teachers to provide the necessary guidance.

In writing these words, I now recognise the damaging consequences of leading such a lifestyle as a child/adolescent, when the quality of mental and emotional development benefits from a supportive and low-stress family environment. What I experienced, rather, was a conflicting combination of nurture and care, mixed with extreme demands for obedience and discipline at almost any cost (certainly at the expense of familial peace and healthy working habits). Furthermore, these extreme experiences at home did not seem to be understood by anyone around me. After my teacher in Boston insisted that I seek counselling to cope with the complicated relationship of me and my mother, I had several interactions with professionals that resulted in disappointment. I felt that they could not fully understand the conflict between the love and gratitude that I felt towards my mother, my deep resentment of her obsessive and volatile behaviour, and a sense of guilt at her sacrifices for both my physical well-being and my musical development. My teacher suggested that I place distance between me and my mother, and avoid discussing matters relating to music or even allowing her to listen to my performances (as her post-performance judgements always seemed to hurt the most). However, I found it impossible to distance myself from the one figure that had thus far been the closest and most significant presence in my life. My friends initially reacted to my experiences with a mixture of confusion, disbelief, and horror, but over time they simply accepted that they would never understand my mother's control over me. Inevitably, I had developed a deep longing and desire for her acceptance and praise, at the risk of my own emotional well-being.

My unhealthy reliance on my mother's recognition of my musical abilities persisted until I moved to London. Once there, I finally managed to put physical distance between my interactions with her, and she was unable to influence my practice beyond my visits home, which occurred only once a year. And yet, I'd still feel the imprint of her controlling behaviour in London, particularly during moments of extreme difficulty and self-doubt.

My aim in detailing this aspect of my upbringing is not to demonise my own mother; nor are my experiences claimed to be representative of the East Asian experience. Rather, by reflecting upon my mother's behaviours, I am attempting to compare and understand them in relation to the discourses that I've explored, and gain the perspective I need to heal (for a second time) from their effects. Returning to Western society's view on East Asian parental discipline, there have been instances where the behaviours of Asian parents have led to severe consequences for the entire family (see Zhao, 2002 for an article on Asian families having been torn apart by child social services; also see Wang, 2015: 186). My own experiences suggest that the complex dynamics between East Asian parents and their children cannot be judged based on Western views on systematic abuse, as often Asian parents' extreme discipline stems from deep concerns for the well-being and happiness of their children's futures.

I have further realised that the societal perpetuation, in the West, of the Asian 'traditional values' of 'reverence for education' and the 'hard-work ethic' (Roth, 1997: 332)⁵⁵ generate a one-dimensional portrayal of East Asian identity, and thus result in a shallow understanding of their experiences. As for East Asian classical musicians, these

⁵⁵ In Kim (2016), it is suggested that the Confucian values of familial piety, harmony, and obedience to elders in East Asia have led to the stifling of Asian children's creativity.

views further reduce their identities to singular characteristics such as ‘superb technicians’. Consequently, this oversimplification by Western society seems to feed their own view of Asian musicians as lacking in artistry and the creativity of ‘an original mind’ (Wang, 2015: 65; also see Walsh, 1983). Sadly, this view may even be perpetrated by East Asian musicians themselves, as they function under the influence of Western society (as mentioned in Chapter One, I was an example of this). Returning again to my own experiences, the discipline of my mother did not necessarily result in my being a ‘superb technician’; rather, she strove to develop my artistry through her own means. Also, her constant demand for obedience did not result in ‘robotic’ behaviour, as I cultivated my skills of argument and critique, rather than followed her instructions. Unfortunately, the dissonance between my own characteristics and those of what I understood to be part of the East Asian identity led me to distance myself from my East Asian background—my mother above all. At the same time, my peculiar upbringing and complex relationship with my parents prevented my full assimilation into Western society. I essentially belonged nowhere.

One behaviour that *is* highly characteristic of East Asian parenting and mentioned often in personal interviews with the parents themselves, is self-sacrifice. Parents speak of their children as having top priority in their lives (Wang 2015). Particularly for the Asian parents of classical musicians in the United States, self-sacrifice is demonstrated in various ways: some have given up successful careers in their home country and, due to barriers involving language or credential evaluation, are forced to work jobs well below their abilities or expertise.⁵⁶ Others have quit their jobs completely in order to dedicate their time and energy to their children’s musical training. Some families are forced to separate as one

⁵⁶ This had been the case for my father, initially, and for my mother, entirely.

parent moves to the new country to support their child while the other stays in Asia, possibly with the remaining children. Due to these changes, families may face financial hardship in their new country, especially when the expense of learning a musical instrument is great. In addition to financial investment, parents who support their children's musical education spend a significant amount of their time and energy driving to and from schools, performances, competitions, as well as supervising their children's practice. Some families are unable to have more children because of the demand of supporting a musical child. While these acts of self-sacrifice may be seen as evidence of the parents' own extreme ambitions and competitiveness, many parents state that the reason for their actions has to do with their children's talent. These parents were told by teachers and other established musicians that their child must be supported to have a life in music; this typically includes moving to a Western country where the child can receive tuition from prestigious institutions and renowned musicians. Some describe having a moral obligation to their children's talent: Wang Wei, a former businesswoman in China, moved to the United States to support her child at the Juilliard Pre-College and believed that she was not a good mother until the move because she was focusing more on her own career (Wang, 2015: 44). Another parent jokingly views music as a 'drug' that, once they and their children have started on the path of pursuing it as a career, feel that they cannot stop devoting their lives to it (Wang, 2015: 46). Regarding my own situation, my mother would often mention the insistence of my first piano teacher that I pursue music, even though she herself was unsure. Thinking that my health was too fragile to continue learning the piano, my mother almost decided that I should stop, but was convinced otherwise, eventually taking my teacher's words as a sign that our family had an *obligation* to continue my musical studies due to my talent. As the years passed and my health recovered, this obligation transformed

into an obsession, spurred on by my mother's own stubbornness to support my path to becoming a successful pianist and to not let my illness lessen my chances for success. Her life subsequently revolved around me and my musical practice, and she sacrificed her own pursuits and interests in order to devote all of her time and energy.

To discuss my mother's behaviour as purely a personal obsession or 'drug' that made it impossible for her to step away from my musical pursuits would, again, be an oversimplification. As I read through the statements of other East Asian parents in Wang (2015) and Yoshihara's (2007) works, I was made to realise the crucial factor that profoundly affected the lives, decisions, and behaviours of so many East Asian parents: the experience of immigration.

Immigration

The interviews and stories of East Asian parents are mostly based in the United States, where it is said that Asians take up instrumental practice in greater numbers than in Europe (Wang, 2015). Whether or not this is true, the majority of cultural and social studies related to East Asian classical musicians discussed in this thesis come from the U.S. It is therefore vital to discuss the topic of immigration in the U.S. in relation to the perceptions and behaviours of Asian parents. In 1965, following the Immigration Act, the United States saw increased immigration of educated professionals from East Asia. These immigrants, most of whom were middle- or upper-middle class in Asia, gained entry based on merit and profession, and subsequently settled into lower-to-middle-class America. When interviewed, these Asian parents carried overseas their ideals about Western classical music as cultural capital. Particularly for Chinese and South Koreans, classical music is associated with the elite and prestige; one immigrant from South Korea described the sense of wealth

she associated with playing the piano during her youth with the fact that ‘my country was not rich enough to educate everyone in piano’. During her college years, she noticed that those who could study classical music were seen as special and prestigious, and that those who listened and enjoyed classical music were distinguished and set apart from ‘regular people’ (Wang, 2015: 54). Those who did not grow up with classical music associated it with edification and sophistication, believing that one develops socially and intellectually by listening and learning to appreciate classical music. One mother from Hong Kong described the popular music that she listened to as a youth as simple and easy to understand, and regretted that she did not have any exposure to classical music. Through her daughter’s training, she learned how to listen for technique and colour, and could ‘feel something’ when she listened to classical music. She also felt that she gained the ‘advantage’ of understanding better; of acquiring more cultural capital herself by investing in her child’s musical education (Wang, 2015: 55).⁵⁷

Although the quest for cultural capital may seem at first glance to support the notion of excessive ambition, further consideration reveals the pressures that Asian parents face in a foreign country where immigrants are a priori deemed unsophisticated due to cultural, social, and economic barriers (Wang, 2015: 55). A common strategy for scaling these barriers involves finding ways to seem intelligent, to be respected, and to be above the rest. Wang cites Lisa Park’s concept of the ‘good immigrant’ to explain this method (Park, 2005:

⁵⁷ It should be noted that, as a footnote, Wang acknowledges that the perception of classical music as the ‘prestigious’ form is not held by all East Asian parents. Some state that both classical and popular music are equally appreciated and valued. It does seem that, for many East Asian parents, their own consumption and enjoyment of popular music (usually from their country of origin) does not interfere with their pursuit of classical music for their children.

6): Park states that this is an adaptation strategy that reveals the pressures on racially diverse immigrants to constantly prove themselves worthy of cultural citizenship, inclusion, and equal rights in the U.S. (Wang, 2015: 57). Additionally, many parents seem aware of the stereotypes and notions about their parenting ways in Western society and use different methods to justify their actions; one method involves projecting the stereotype onto others. South Korean and Chinese parents at Juilliard Pre-College generally form their own separate groups and find it difficult to mingle between groups; this may largely be due to language barriers, but the result is that each group seeks to seem more balanced and less status-driven than the other. Some Korean parents speak of the Chinese as being more competitive, or having a stronger culture, and therefore able to be more disciplined and fight for what they want. One Chinese parent seemed amazed at the discipline of Korean children, stating, 'They don't have TV, play time, practice 7-8 hours a day...'; another suggests that Korean parents are merely interested in cultural status and producing the 'next Sarah Chang' (Wang, 2015: 49). Another method is to criticise the behaviour of other Asian families by stating the dangers of too much sacrifice and pressure to succeed. One parent claims that other Asian parents don't like her because she tells them that it is not worth it to push their children so much; to break up families and put too much pressure upon the children to be performers and soloists. Interestingly, she also suggests that there may be other issues such as marital strife or discord that separates families, hinting that the sacrifices that some parents make are not purely for the sake of their children's musical training.

Despite the supposed differences between Chinese and Korean parental groups, East Asian parents' perspectives on classical music training are often united in one respect: Western classical music is viewed as more of an Eastern, rather than a Western, pursuit. As

Wang (2015: 61) suggests, East Asian parents, in an attempt to find their own, specialised space within American society, believe that the effort necessarily involved in classical music training expresses Asian values: sacrifice for the good of their children, appreciation for hard work, and so on. They tend to view Western parents as individualistic and self-centred, unwilling to sacrifice, preferring to seek out easy paths and quick wins in life (Wang, 2015: 50). By pointing to these discrepancies between 'Asian' and 'American' views on parenting, Asian parents seem to accept or at least recognise their exclusion from American society, and attempt to defend their own space by claiming Western classical music as their own. It is important to acknowledge that Asian immigrant parents function under the awareness of racism, social and economic barriers, downward mobility, and language discrimination in their new country. Perhaps, rather than their own social aspiration, it is in order to protect their children (even those who are born in a Western country) against such pressures that Asian parents impose control and discipline on their offspring. Additionally, they navigate the difficulties that come with the misalignment of their parenting with mainstream Western standards, such as losing their children to the state, or causing damage to their children by pushing too much while the latter are being surrounded by contrasting views and demands. Despite the critical perception from Western society, and the anxiety of being aware of potential conflict and damage to their children, Asian parents continue to dedicate themselves towards supporting their children's musical training. This is due to their belief that, in the field of Western classical music, talent and skill can allow their children to overcome barriers, to 'be somebody', so that they can get the 'right treatment' (Wang, 2015: 60).

I have come to view my own mother's desire to ensure my musical success as stemming from the fears, insecurities and conflicts that she experienced both in navigating a

vastly different and challenging new world and in coping with her own child's severe illness. Much of the statements mentioned in this final section resonated deeply with me; I have always been aware of, having observed first-hand, the disparity between the dreams of a better life that determined my parents' move to the United States and the reality of the several years that followed this decision. My father, having held a post-doctoral research position in China, was forced to take on manual labour in order to support his family. Although his status and position eventually improved, my mother suffered the opposite. Isolated from society by her decision to stay at home, she devoted her life to my own while her skills in English communication failed to go beyond an elementary level. Along with the language barrier, she was further burdened by financial insecurity and the constant worry over the health and future of her child, aspects that she seemed to view as her failings as a mother. It now comes as no surprise to me that music became her refuge; along with the above-mentioned associations with status and sophistication, classical music became the only outlet for my mother to express the abilities and the creativity that she possesses. Although her forceful authority over me and my music-making may have threatened the development of my own authoritative voice, I have gained the perspective necessary to grapple with this issue.

By weaving my own background into this discussion of East Asians as a 'model' minority, I have sought to provide more nuance to the discourses that surround the East Asian experience in society, and particularly in Western classical music. However, a further—and even more crucial—aim has been to give voice to my difficult bicultural experiences, in order to better understand the connection between my research concerns and myself. My feelings of alienation and difference developed from growing up within two vastly

contradicting cultural environments: my life at home was an alternate reality to that of the outside world, and the society in which I functioned lacked the necessary understanding to provide me with the support I needed. I managed to cope with my alienation by focusing on improving my skills and knowledge in music, and by disassociating myself from either an Eastern or a Western cultural identity. However, the combination of increasing pressures and demands to achieve from within the competitive environment of the conservatoire; the existence of subtle, racialised discourses in society; and certain orthodox views on classical music as a Western art, entrenched in the histories, cultures, and traditions of the Western world; have resulted in further insecurities and doubts about my position within classical music—that is, within the very refuge I had made for myself. This leads to the final section of this chapter, in which I address the complex and conflicting concept of authenticity within Western classical music and its particular effects upon East Asian musicians. This includes confronting the tensions between staying true to the composer’s intentions while seeking an authoritative voice as a performer.

5.5 Authenticity in Western Classical Music

It is inevitable that, when discussing Asian identity in Western classical music, or simply Western classical music in general, the complex issue of authenticity arises. Authenticity plays a significant role in discussions of the nature of all artistic practice within an ever-changing, global society. Generally defined, the concept refers to origins and tradition, doing things the way they have always been done, and being faithful to the original.⁵⁸ However, in today’s complex cultural climate, this term has been contested in several ways,

⁵⁸ See the *Oxford English Dictionary* definition.

particularly in the arts and academia. Firstly, the association of authenticity with something fixed and unchangeable has been challenged by cultural commentators; Walter Benjamin argued that authenticity, or rather 'tradition', was something 'thoroughly alive and extraordinarily changeable' (Benjamin, 1969: 174). He believed that the authenticity of an object involves its '*full* historical testimony, that is, the entire range of contexts it has passed through, not just in its singular, documentary testimony of its origins' (in McCole, 1993: 6). Others have supported this view by revealing that what are supposed to be the ancient traditions of a culture or a people are often the result of relatively recent historical or political developments, many of which are symptoms of modernity (Wagner, 1981). Deconstructionism and poststructuralism have challenged the existence of 'legitimate' knowledge, or a 'pure' object of enquiry that exists and functions independently of outside forces, including the presence and gaze of the observer (Clifford, 1988). Postcolonial studies have revealed that colonialism has created cultural hybrids of both the colonised and the colonisers, and globalisation studies have shown that cultures are spread and practised in multidimensional ways across geographical borders (Yoshihara, 2007: 190). The authority of past cultural artefacts or practices is in fact often derived from the present (see Taruskin, 1988 on authenticity and historical performance). However, as Yoshihara (2007) reminds us, there are those who still cling on to a fixed idea of authenticity and tradition, those who view their connections with culture as being diluted and threatened by products of modernity and postmodernity (global capitalism, or mass culture). These people seek a 'culturally and historically-rooted, socially-located self' (Yoshihara, 2007: 190).

As it so happens, the field of Western classical music has largely grappled with the imposing concept of authenticity due to its rootedness in tradition and history. As much of its canonic repertoire is from the distant past, there is a predominant focus on the

composers and the compositions of the past, rather than the performers or the performances of the present. From the mid-19th century until relatively recently, the orthodoxy of historical musicology has been to concentrate on the composer and his (or her, but not by coincidence usually his) musical, cultural, and historical context, rather than that of the performer. Even in the concert hall, the performer's task has, during this period, been reduced to meticulous study and rehearsal of the composer's music with the purpose of providing a more or less neutral connection between the past and the present (i.e. listeners). This discourse is widely reinforced within music pedagogy, where significant demand is placed upon the student to have an accurate understanding of musical styles and periods (and their conventions); particularly for instrumentalists, there is pressure to adhere faithfully to the score, as the repository of the composer's genius.

The notion that performing instrumentalists gain legitimacy by being faithful to the score, which is seen as an 'authentic' object that is historically, culturally, and musically bound to the person and circumstances of its creation, is beginning to be challenged. Doğantan-Dack's (2012) work sets out a parallel discourse that places the focus on the performer's musical individuality and expression. According to Doğantan-Dack, the most significant and consequential contribution to contemporary discourse on Western classical music performance is the development of the notion of 'the music' as an entity that is differentiated from that of the 'the score', 'the work', and the performer. This complex relationship has led to the conception of the performer as mere 'interpreter' (Doğantan-Dack, 2012: 8). The complexity involved in the defining of 'the music', 'the score' and 'the work' created two distinct, yet parallel discourses of performance which are still relevant today in relation to authenticity.

The first discourse, developed in the 19th century, views ‘the score’, ‘the work’, and ‘the music’ as a single entity to which both the performer and the performance must be subservient (see Goehr, 1992). Goehr (1992: 231) reasons that, due to the successful development of a ‘complete and adequate’ notational system, ‘the score’ was deemed entirely sufficient at representing ‘the work’. Additionally, Goehr uses the terms ‘the work’ and ‘the music’ interchangeably when she states that, ‘the perfect performance of music is the perfect performance of the work’ (Goehr, 2002: 141; quoted in Doğantan-Dack, 2012: 10). The existence of the ‘work-concept’ implies that the score is a direct, complete and inviolable representation of the music imagined and created by the composer. Increasing numbers of influential musical figures (theorists and composers) held this view during the 19th century and beyond, including E.T.A. Hoffmann (Charlton, 1989), who believed that in the performance of a composer’s work, the performer’s personality must not in any way interfere. Composers such as Hindemith, Stravinsky, and Schoenberg continued vigorously to uphold this view into the 20th century (see footnotes in Doğantan-Dack, 2012: 10).

The second, parallel view argues that ‘the music’ is not found in the notated score, but rather in the act of performance. Furthermore, a performer cannot logically limit her or his role to replicating the work of the composer, since the act of performance unavoidably involves the creative influence of the performer who is committing the act. Franz Liszt famously argued that the performer is ‘not merely a mason who, chisel in hand, faithfully and conscientiously whittles stone after the design of an architect ... He creates as the composer himself created’ (Sachs, 1982; quoted in Doğantan-Dack, 2012: 8). Likewise, Liszt’s admirer Anton Rubinstein believed that it was impossible for an *objective* interpretation of music to exist; stating that, although it is the ‘law and duty’ of the interpreter (i.e. performer) to ‘do justice to the *object* (i.e. composition)’, they inevitably

must do it their own way, 'i.e. *subjectively*' (trans. Morgan, 189?; quoted in Doğantan-Dack, 2012: 9). According to Rubenstein, the concepts of the composition and the interpreted music are separated. The implications of this second view have resulted in the conception of the 'music-behind-the-score', which focuses on music as a 'sounding phenomenon' that involves expressive details executed by the performer that are not (and cannot be) specified in the score (Doğantan-Dack, 2012: 11). This leads to a heightened focus on the performer and the performance, as well as the development of a contemporary discourse in contrast to the orthodox, composer-centred view mentioned earlier (see Clarke and Cook, 2004 for an overview of landmarks in Performance Studies; also see Cook, 2013).

At some point in our education, all musicians trained in Western classical music are exposed to the tensions created by two discourses: one, the score is the ultimate object of authority, and the performer's role is to obey the text; and two, individual interpretation and expression are essential for performance and our principal *raison d'être*. Eventually, we find a more or less happy medium between the two, balancing sometimes precariously in between and constantly adjusting when we fall too far to one side or the other. For some musicians, however, there can exist conflict between two further, opposing discourses: one, Western classical music is universal and above any racial, political, or cultural influence; and two, Western classical music is interpreted by the performer, therefore performances, due to the racial and cultural plurality of performers in current times, are inevitably racialised and culturally specific. The next section details various ways in which East Asian musicians view themselves within one discourse or the other, and how they deal with the existence of both.

Voices: Interviews with East Asian Classical Musicians

Categories of the Universalist view

Returning to Yoshihara (2007), the interviewed Asian musicians fell into differing categories of opinion on authenticity and one's racial and cultural identity. The first category involved those who did not want to be associated with views of racial and cultural influence in music-making at all; these musicians found offensive, or thought it irrelevant, to take part in the discussions on authenticity, at least as it pertains to race and culture, and decided against contributing to the interviews. Those in the second category of musicians agreed to be interviewed, but openly dismissed the question of authenticity as irrelevant. One Asian violinist mentioned an instance when a French musician asked her if she found it difficult to play Western music, because when he played German music, he had a hard time understanding it. The violinist found the question appalling and offensive, stating that 'it's almost impossible to feel your nationality or race in music. It has never, ever crossed my mind.' As for the French musician, the violinist believed that 'this guy is probably never going to have a career' because 'these days, when we are able to travel across the world within a day, you can go and experience just as much what other people are experiencing, taste their food, adapt to their lifestyle' (Yoshihara, 2007: 199). It is interesting to note that, for this particular Asian violinist, music could be better understood if one exposes oneself to the experiences of another culture. Another interviewee, a pianist, pointed out that music from another culture can be accessed and understood through the internet, where one has a neutral kind of access to the world. She believed that it is up to the performer to 'capture, absorb, and express the essence of any music' and that if one cannot convincingly play a piece of music from another culture, that means that one has no talent (Yoshihara, 2007:

200). A third category of opinions involves a universalist view of Western classical music, where no particular culture or group of people 'owns' classical music. One example is that of a double bassist, who believed that while Western classical music may have originated in the West, it is no longer 'a form of European ethnic music', just as Christianity is no longer solely a 'Western' religion due to having many different elements added to it over time (Yoshihara, 2007: 201). Classical music, according to this interviewee, has been elevated to a universal level. Furthermore, he claimed that the act of playing music is simply 'a very concrete act of making sound through physical movements'. He did not believe that there is anything more to playing Western classical music than knowing 'how to hold the string or how to move the bow and how to make things sound big or small'; as long as one acquires the proper technique, one can learn how to do things that are sometimes claimed to be innate, such as characteristic rhythms (Yoshihara, 2007: 202).

All three of these categories deny the existence of a culturally specific essence within Western classical music. They offer several ways in which people from any culture can convincingly play and 'own' classical music: one, expose yourself to the experiences, food, and lifestyle of a culture (unspecified but presumably Western); two, make use of the free-flow of information brought on by globalisation to expose yourself to the music; and three, acquire the proper technique which allows you to succeed in the physical act of playing music. Although these simple methods to music-making seem to shine a positive light on East Asian musicians' position within the world of Western classical music, complexity and conflict can be teased out from behind the words of the interviewees. The first brushed off the struggles of the French musician, possibly because she was offended by his question, claiming that 'he will never have a career' if he thought that way. However, through my own experiences and hearing about those of other musicians, I believe that such an awareness of

vulnerabilities is a sign of deeply reflective and critical thinking, which leads to further development, strength, and awareness about what it is that we do as musicians. This, in turn, serves the music we play, as well as the knowledge that can be passed onto others.

The second interviewee used an imposing word that deserves to be further unpacked, but is unfortunately beyond the scope of this thesis: talent. This association between the inability to play music from another culture convincingly and a lack of talent may lead to defeatist views about one's own abilities, thereby stifling one's desire for improvement and development in music-making. Rather than relying on the existence of an imprecisely-defined level of aptitude (passive state), it is surely more productive to focus on the aim of improvement and discover the means that will aid in one's development (active state). The third interviewee did mention the importance of action in terms of acquiring proper technique, yet his view of music-making was limited to purely physical actions. As a pianist, I am constantly aware of the complex relationship between the mental and the physical acts involved in playing an instrument that is fundamentally limited in its ability to sustain sound or adjust the tuning of pitches. In my experience, a significant portion of music-making on the piano involves activating one's imagination and 'inner hearing', rather than relying solely on the 'proper technique', which can be variable depending on the performer and her or his physique and attributes. Of course, the interviewee may have been referring to the most basic level of music-making, which involves the accurate performance of the notation: articulation, tempo, and dynamic markings. To this extent, it may be true that 'anyone can do those things'; however, challenges arise when one aims for further meaning in their performance, seeking to progress to true artistry.

Categories of the Particularist view

The interviewees mentioned above believed that authenticity, in the narrow sense of an essence possessed only by certain racial and cultural groups, does not exist in Western classical music. They supported a universalist outlook that denies that any particular culture holds ownership of Western classical music; they also viewed music-making to be something that can in principle be learned, understood, and practised by all. However, there were also those interviewed who did believe in some form of racial and cultural essence to music-making, revealing a particularist view that takes into consideration the specific contexts of various musical sensibilities, as well as the involvement and impact of culture and identity upon the practice of classical music.

There are categories of particularist responses, too. First, there were those who believed that rhythm is an obvious marker of one's 'innate feeling' for music, and that it is 'in the blood' of those who share the national or ethnic culture of the composer. One Asian violinist stated that when she heard a Bulgarian musician play a piece derived from Bulgarian folk music, there was a 'natural rhythm' to the performance that she felt incapable of capturing. Likewise, a Japanese pianist commented that 'What we have in our blood is inherently in us [...] The sense of rhythm that's culturally specific becomes most clear when we play ethnic dance music, like polonaise, Hungarian dance music, things like that' (Yoshihara, 2007: 202).⁵⁹ Second, some interviewees mentioned an inexplicable 'essence' that Western musicians possess 'in their blood', one that Asian musicians struggle

⁵⁹ Here, I refer to Chapter Two, where I cited studies conducted on the performance of rhythm with Japanese musicians, based on the notion that Japanese musicians play certain rhythms in Western classical music differently from Western musicians (Sadakata and Ohgushi, 2004; Ohgushi, 1999, 2002).

to understand or possess themselves. A Japanese trombonist commented: 'when Europeans play music like Mahler, Mozart, or Beethoven, it's hard enough for me just to follow along... I do feel that those people who have it in their blood relate to it differently'. He continued to say that he viewed himself as 'a Japanese musician who has *studied* Western music' and believed there are limitations to that position. Also, although cultural boundaries can be crossed, he believed that it would take 'a whole lifetime. You keep working hard, with open ears and open mind, and maybe at the end of your career you will reach that goal'.

(Yoshihara, 2007: 203). Another Asian musician commented on the 'minute details' that distinguish 'the essence' from superficiality. She explained that 'for example, a Japanese playing Mahler sometimes would do things that a German musician would never do. It is a really minute thing, but still definite and specific, and it makes a difference'. This particular musician believed that a listener would be able to tell the difference between the performances of an Asian and a Western musician because 'there is an unspoken agreement between the performer and the listener that has been made historically and culturally'

(Yoshihara, 2007: 203). It is important to notice the lack of specificity in this musician's comments about the difference, and the 'unspoken agreement'. This is representative of the complexity that continues to create tension and conflict within Asian classical musicians. The qualities or 'unspoken rules' that make a performance 'authentic' or 'inauthentic' remain largely undefinable. Some musicians, such as the ones in the third category, believe that culture is involved in these differences: one musician stated that her lack of knowledge about Christianity made it difficult to understand religious pieces; another mentioned the difference between her sensitivities to major and minor tonalities and those of other races and cultures; a third interviewee, echoing some of the ideas that I have been discussing here, thought that Asian musicians are inhibited with regards to musical expression because

of how children are taught to behave. He stated that in Asian culture, it is considered inappropriate to show too much of one's emotions, and that due to the restrictions placed on expressing certain ideas or a wide range of emotions, Asian musicians tend to 'go inward' and require 'some extra element' to let go, expressively.

At this point, it is necessary to locate myself within a 'particularistic' position that involves the central topic of my thesis: language. Some Asian musicians who occupy this category contradict the statement that 'classical music is a universal language' by mentioning instances when the native language of a performer seem to influence their musical expression. One pianist recalled hearing a student from Hong Kong perform, with a degree of struggle, Chopin's third piano sonata. She explained that the student had difficulty with phrasing, sounding, as her friend later pointed out, like she was 'speaking' Cantonese in her performance. The inflections of Cantonese, according to the pianist, are completely different from the inflections of French,⁶⁰ and this affected the student's phrasing. At this point, the pianist realised that learning Western classical music, as someone whose native language is different from that of the composer, is similar to learning a foreign language: one can 'hit all of the right notes', but still fail to sound 'idiomatic'. She also commented that Asians who are raised in Europe or in the United States (and have experience with the languages of Western classical composers) might have a different relationship to classical music from that of those raised in Asia (where, as I have been discussing, the morphology of the languages is vastly different). This suggests that linguistic identity is separate from racial or ethnic identity; it also implies that one can come to 'acquire' musical expression, just as

⁶⁰ Interestingly, this pianist views Chopin's language as French rather than Polish, his native tongue; there are many musicians who would disagree with this.

one learns a new language. In the words of the Asian pianist, 'Understanding music is like understanding a verbal language, an acquired skill' (Yoshihara, 2007: 206). Clearly, this challenges the perception of near-immutability regarding the 'essence' of Western classical music that some of the interviewees hold. When musicians associate their music-making with language and speech, in other words, they seem to find themselves in possession of more tools for mastering the large variety of music that they must convincingly perform. The world-renowned violinist Midori provided an anecdote about her experiences with Bartók's music and Hungarian speech: 'For instance, Bartok. His music is closely related to Hungarian diction and all that. Sometimes I do think about it... I always learn by the words, and sometimes I think it's by the way the language sounds, the diction, the breathing, or the nasalness of that language, that I try to incorporate in my music-making when I play it on the violin. So in a way it's so closely related to language' (Wang, 2015: 69).

It should be made clear that the majority of interviewees in Yoshihara's study held beliefs that span several of the categories mentioned in this section. Although some gave elaborate explanations of the 'essence' in classical music, they concluded that it is not their goal to 'learn the essence of Western music and to acquire its ultimate truth'. Rather, their goal is to make the music their own. Many also spoke of the changes that are impinging on the way classical music is heard and performed in today's globalised world. One violinist stated that, 'as more and more people who don't follow the conventions come to play Western music, those conventions themselves will eventually change' (Yoshihara, 2007: 208). An Asian harpist commented that, '[t]oday, both the musicians and their audiences travel all over the world, and the idea of authenticity itself is becoming meaningless...' (Yoshihara, 2007: 209). These Asian musicians recognised certain 'essences' in Western classical music, as well as the subtle differences in musical expression between Asian and

Western musicians, but remained confident in their own musical voices and identities. Yet for those who have been strongly affected by the awareness of subtle differences, it is not necessarily beneficial to internalise the discourse of universality, which ignores specificity of culture, history, and language. Rather than brushing off certain doubts and struggles as signs of lack of talent, or unsuitability for an international career, these doubts can be recognised and analysed on both an individual and a social level, with the purpose of strengthening self-identity and achieving further artistic integrity in music-making.

This point leads to the final chapter of my thesis, in which I conclude the several years of research and musical practice that has led to the realisation that I hold multiple identities, formed through my experiences as a Chinese-born, Mandarin- and English-speaking classical pianist who was raised and educated in the U.S. and continued studies in both musical performance and research in the U.K. My sensitivity to, and subsequent interest in, the conflicts and tensions that arise from differing backgrounds and experiences have equipped me, I now find, with the necessary skills to flourish as a chamber musician. With the final musical performance that accompanies this thesis, I aim to reflect my various identities through my musical voice.

Chapter Six: Conclusion

6.1 Summary

There is something about the Berg sonata, I have come to realise, that distinguishes it from all the other pieces I have learned so far. Perhaps it is the profound expressivity of the music: confusing yet captivating in its complexity, in part due to its unique musical language (and here I am using the term in the traditional musicological sense, that is, in reference to compositional elements of harmony, rhythm, motivic variation, counterpoint, voice-leading and so on), and further heightened by a remarkable mixture of lyrical beauty and ‘tragic poignancy’ (Carner, 1983: 114). More than this, however, I am convinced that it is the impact it has had on my understanding of myself that sets this piece apart from others. This music has forced me to look within myself, to question what I know and feel, and to engage with how I express both my knowledge and instincts through performance.

In developing a relationship with this piece I went through an identity crisis. I found myself deficient in many respects: my sense of who I was (musically, culturally, and personally), my grasp of the kind of musical knowledge that bridges the theoretical and the practical. These things are necessary for deciphering music like a language and communicating with it like speech. My crisis made me aware of how little authority I had over my own music-making, and, crucially, how little voice I had as a performer.

Slowly, through research, I began to understand the multifaceted nature of speech and musical performance, and how that nature is bound up with identity. Through experimentation, I came to understand more about the variability of individuals’ expression in both speech and music. By clarifying and disaggregating the different systems and parameters of speech production and musical performance, I realised that the traditional

ways in which musicians attribute expressiveness to an individual's music-making may be, variously, oversimplified, overly mystified, and/or misguided: the common associations of expression with talent and inherent musicality make expressivity seem unteachable and unsusceptible to development. Furthermore, differences in expressivity are often judged as either right or wrong. Too much of that judgement is typically located with obscure notions of authority and authenticity (tradition, style and so on) and too little with acknowledgement of the complex and interacting factors involved in each expressive decision.

From research-led reflections on my practice and performance of both solo and chamber music to practice-as-research in the form of some uniquely stimulating experiences during lessons and rehearsals with other musicians, I gained more knowledge of the connections between speech and music-making, as well as more confidence in my position as a performer-researcher. I also recognised my particular strengths as a chamber musician, and made a conscious decision to develop my performing career within that discipline. Finally, my cultural survey of the East-Asian experience in Western classical music has united my own personal experiences (as a child of East Asian parents, struggling between two distinct and at times conflicting cultures) with those of other musicians from similar backgrounds. I felt recognised in the words and experiences of others, and, in turn, recognised how little I myself knew about the complex history of Western classical music in East Asia. I also confronted the complicated relationships that we, and our East Asian parents, have with Western classical music and with one another.

6.2 Embodied Awareness

The sensation of performing feels different from what I remember from four years ago. Of course, by this I don't mean to idealise my musical growth, as if it has resulted solely from my research journey. There are obviously some straightforward explanations for the changes that I notice in myself: I am older now, and with age comes experience, as well as a sense of authority that develops naturally with the passing of time. I have continued to have lessons, perform, learn new repertoire, and revise older pieces through more informed and improved processes over the course of four years, while feeling more stable in both my personal relationships and my physical environment. These, and other subtle aspects of my everyday life, have undeniably contributed to the increase in mental and physical confidence that I experience now when I play the Berg. However, I do believe that my systematic investigation of personal concerns, prompted initially by musical ones, has brought me full circle. My journey has been significant to my musical development.

To provide a specific example: when I play the Berg sonata now, I am more aware of the music's tensions and resolutions in a physical, or rather embodied, sense than before.⁶¹ This awareness, which extends, to a greater or lesser extent, to all of the music that I currently play, is not the same as simply knowing how the compositional elements fit together. Managing this distinction is also not especially rare: to quote Mannes in the Foreword of Salzer (1952), there are many musicians who experience the 'seemingly unbridgeable gap between their theoretical studies and the living experience of music itself' (Salzer, 1952: vii). But, as I have mentioned above, part of my self-described identity crisis

⁶¹ For research on the expressive functions of physical movement in musical performance, including embodied musical structure in piano performance, see Buck et al., 2013.

involved the confrontation of my inadequacies in both theoretical knowledge of music and its integration into practice.⁶²

I currently experience moments during performance in which my listening to the music aligns with certain physical responses and gestures that I can only describe as instinctive. I can describe these moments as extremely pleasurable in the sense that I feel fully in control of my expressive decisions as they emanate from within myself, rather than from external influences such as the performances of others, or others' suggestions of how I should execute the decisions that they have made for me. A brief and preliminary search in current scholarship reveals a growing interest in the topic of the body and its connection with musical knowledge (see Schiavio et al., 2019 for an overview), encouraging me to explore further and seek to understand these pleasurable moments in performance. Importantly, this feeling of pleasure brings a sense of control, or power, along with freedom. I rarely doubt the correctness of my expressive decisions anymore (of course, this is an ongoing process, with the occasional relapse), and am currently satisfied with the knowledge that, if someone were to question my decisions and suggest another way to more convincingly express the music, I can face this as a learning experience, or an opportunity to engage, rather than a direct attack on my abilities as a musician. This, for me, is a major development.

⁶² For an interesting model of musical experience as 'encounters', in which aspects associated with the varying experiences of music are divided into three 'experiential realms', 'pragmatic', 'esoteric', and 'abstract', see Moran (2014).

6.3 Revising and Reforming

The Fauré Sonata for Violin and Piano in A major is a piece that I first learned around six years ago. The music seems to swirl with colourful passion and a sincere outpouring of expression; both the experiences of performing and listening to this piece are emotionally rewarding and immensely satisfying. Returning to this again, after so many years and such a long research journey, I now face a number of new challenges. My performance six years ago was during the final Master's recital of a dear and close friend in Boston; my strongest memory of the performance is my association of the music with a final farewell. The violinist and I were both graduating, about to follow separate paths, and we were not sure when we would ever see each other again, let alone play together. The emotional associations of friendship and separation that I formed with the music, combined with the aesthetics of sensibility and sincerity inherent in Fauré's work (see Caballero, 2001), came through in performance and were palpable for the audience. However, at that point in my musical education, I did not view the specificities within the music as part of a 'language', as I do now; my process of understanding the piece involved playing through the music many times and carving out an interpretation as I adjusted to what I heard as 'beautiful' and 'musical', or not. This process relied on my instincts as an appreciative listener of music, rather than an informed and fluent 'speaker'. Although the performance was a successful and enjoyable experience, I realised as I started my revision that I had not achieved a very deep understanding of the music, and thus it now felt new and, significantly, *foreign*. I found myself reverting to remnants of the expressive decisions that my body had absorbed (and still remembered) from all those years ago, but it was like an out-of-body experience. I had to fight against certain automatic movements and tendencies, conscious of the fact that I was not actively making these decisions with the instinctive pleasure and power that I

earlier described. Finally, the solution for overcoming these feelings was to imagine that I had never seen this music before. I studied it like a piece of text, familiarising myself with the composer's musical language and decisions (both compositional and expressive), as well as the sounds of French speech through clips on YouTube (I even took inspiration from listening to the casual conversations of my French friends); I also supplemented my study with intense practice sessions of 'inner hearing' (which involves techniques such as the physical singing one textual voice, for example the soprano line, while hearing the bass line in my head), as well as scholarship on Fauré's unique harmonic language and compositional background (see Greer, 1991). These various methods are just some of the ways in which I sought to achieve a deeper understanding of the music. They are not meant to make up an exhaustive list or define a learning process; rather, they are evidence of the agency I now feel I possess to develop as a performer independently.

6.4 Revitalising Connections

My practice of the Schumann Piano Quartet directly confronts the consequences of a traumatic incident that occurred many years ago, one that had a significant detrimental effect upon the development of my musical identity. I must preface my description of the incident, which involved my mother's post-performance critique, with the observation that if at the time, I had possessed the knowledge, experience, and perspective that I have now, my musical identity may not have suffered as much. But perhaps I would also not have arrived at my current position of relative empowerment, if not for that and other similar experiences.

I initially learned the quartet during a chamber music course in 2010. The supportive environment there, combined with the open appreciation of my music-making by my

teachers and fellow students, allowed my emotional connection with and appreciation of the piece to flourish. The faculty focused on heightening emotional involvement with music and communication between players in the ensemble. At the time, I thought of these elements as my strengths; I had come to form my musical identity around being an expressive and communicative performer, rather than one who was technically brilliant but musically limited (the common prejudice against East Asian classical performers). The two weeks of intensive coaching, rehearsals, and social interactions with other highly gifted students led to a final performance of the Schumann Quartet in front of the entire campus as well as the families/friends of the participants. My parents and brother were in attendance, and I had an enjoyable experience during the performance. Afterwards, I was congratulated, as per post-performance etiquette, by the faculty and other students. My parents came to do the same, and I immediately sensed something wrong in my mother's behaviour. Her smile seemed forced, and she rushed to mention that there were things in my performance that she needed to tell me about, that I needed to notice and change. My buzz of energy and satisfaction from having just performed abruptly plummeted at these words. We returned to my room in the student dormitory and I listened to my mother's observations as I changed out of my concert dress. I was moving around too much, she said, and although she was aware that this is usually a sign of engagement with both the music as well as the other players, what I was doing seemed superficial and forced. For her, my performance in general did not make much sense, and the reasons for this were lost to me because I experienced what I can only understand as an emotional breakdown. I began to argue with my mother, explaining my intentions and musical decisions, and the argument escalated into shouting. In the back of my mind, I knew that the other students in the dorm

could hear us, and that, although this was a normal occurrence in our own household, it was incredibly shocking to others who did not have such experiences.

Although I was forceful in my argument with my mother, her words about my superficial musicality struck a chord: I had, of course, similar concerns deep within myself. After that incident, I decided that I was in need of change, that I could no longer assume the identity of a musical and emotive performer. Subsequently, I began to focus intently on my technique—but also on my habit of relying on physical movement to express emotion and engagement. Although my technical approach to playing improved significantly (through the guidance of a second teacher), my musical expressivity suffered (my first teacher was deeply concerned at my seemingly abrupt lack of emotional engagement and ‘musicality’). I was unfazed by this, convinced that I had to strip away the emotional attachment that I had with music in order to arrive at a pure and serious approach.

Returning to the Schumann after such a period of uncertainty and adjustment has been revelatory. As with the Fauré, I began the revision process by studying and analysing the text, forming a harmonic and tonal map of the piece in order to understand its structure. As I played through the music, I was initially uncomfortable with certain structural, rhetorical, and narrative elements. For example, the juxtapositions within the first movement of the simple, hymn-like introduction, the extroverted principal theme (which interestingly takes its material from the introduction, but in a completely contrasting character), and the swirling figurations that follow felt strange and almost silly due to their sudden contrasts. Further along the movement, the music pushes and pulls in ways that felt contrary to my own nature: rather than having control and stability, it felt sudden, impulsive, and exuberant. The process of relearning this music was initially frustrating as the

more I familiarised myself with the musical idioms and rhetorical gestures,⁶³ the more I found them inconsistent with my own temperament. However, as I progressed in my revision, I became more and more fascinated with Schumann's creativity and emotionality. By practising this music, I was confronting directly the artistic decisions of someone whose nature had obviously been wildly different from my own. In the end, I was able to transcend my discomfort and allow this piece to revitalise the emotional connection with music that I had once lost. During rehearsals with the rest of the ensemble, I felt both a sense of power and an emotional surrender as I wove my textures into and around theirs, at certain times pushing forcefully forward and, at others, pulling the music back.

I will detail a final experience to end this chapter, and the entire thesis. It is a moment during a recent concert performance, in which I became fully aware of the expressive and communicative power that I possess.

6.5 An Encore: Schumann Piano Quartet, *Andante cantabile*

'What should we do? They're still clapping' 'We should play an encore' 'But what? The scherzo? Without repeats' 'No, the slow movement, the Andante' 'Okay, yeah, the slow movement' 'No repeat'.

I made the decision in favour of the slow movement while we bowed to the audience for a fourth time. Our concert, initially stressful due to travel complications and limited rehearsal time, had been transformed into one of the most enjoyable performances that I've experienced this year. As the tensions and worries left our bodies following the completion of

⁶³ These elements, according to Gertsch (2013), make up the musical 'persona' in Schumann's Piano Quartet; the term was first conceived in Cone (1974).

the performances of three piano quartets, we made our way back to our seats for the encore. After a brief smile to one another, we launched into the dissonance of the diminished-seventh chord that opens the third movement of Schumann's Piano Quartet, leaning upon it much longer than during the initial performance. As the music flowed between and around us, I allowed myself to exaggerate my decisions of timing rubato, of variations in stress and colour, feeling—in the most present physical way possible—the complete acceptance of my musical companions. When the violinist took over the melody of the cellist, restating it with a different yet equally moving voice, I could sense a change in how she was characterising the line from the previous performance: it was more intimate and vulnerable. Eventually, we arrived just before the moment when the violinist would always pull back the tempo for a brief and delicate instant, to express a poignant slide over the dark and hopeless quality of the G minor harmony. But something was different this time; I felt her intention change—she seemed to want to move through. In that split second, I considered her decision; probably she wanted to contrast with the more intimate character of her restatement of the melody. However, I felt that this moment needed even more time than before, to express more bitter sadness among the sublimely intimate mood we had cast. Instinctively, I slowed my repeated chords which were supporting the violinist's line, pulling back the tempo to a point that shocked even me, and she answered with the most delicate and soft colour that I had heard from her during our few days together. It was immensely moving and also astonishing that my intention had been communicated so clearly to her, and that she had listened and responded in agreement. After the concert was over, we light-heartedly discussed this moment, with her admitting that she had wanted to do something different, but was convinced by me in the end. The rest of the ensemble also commented with appreciation at the effect of our eventually united decision. I made a

mental note always to remember this. I had possessed enough authority to influence another—highly accomplished and experienced—musician’s decision during performance. My musical voice had been heard, understood, and accepted.

At the end of this journey, I have come to recognise the multiplicity of my identities, each one bound in different ways to the varying experiences and complex relationships that have coloured my life. In trying to find my musical voice, I realised that my musical identity, which originally had been the only identity I accepted and deemed worth developing, is one of many: I am also Chinese-American, raised by immigrant parents who struggled and overcame immense difficulties, both in their own lives and in mine; I am also someone who has experienced the crippling pain and devastation of a life-threatening illness, at an age when many experience youthful innocence and carefree exploration; I am also a researcher who has conducted rigorous and systematic investigation, contributing new knowledge, however small, to scholarship in multiple disciplines. Finally, I am a performer with a voice that, inevitably and unashamedly, expresses all of myself through music.

Appendix A: Participant Consent Form

Research Participant Consent

Title of project: Finding My Voice: An Interdisciplinary and Multi-Methodological Investigation into the Relationship between Performers' Speech and Musical Expression

Study approved by School Research Ethics Committee:

Thank you for agreeing to take part in this research. The person organising the research must explain the project to you and you should have read any accompanying information sheet before you complete this form.

- ☐ If you have any questions arising from the Information Sheet or explanation already given to you, please ask the researcher before you decide whether to participate. You will be given a copy of this Consent Form to keep and refer to at any time.
- ☐ I understand that if I decide at any time during the research that I no longer wish to participate in this project, I can notify the researchers involved and withdraw from it immediately without giving any reason.
- ☐ I consent to the processing of my personal information for the purposes explained to me. I understand that such information will be treated in accordance with the terms of the Data Protection Act 1998.
- ☐ I agree that the research team may use my data for future research and understand that any such use of identifiable data would be reviewed and approved by a research ethics committee. (In such cases, as with this project, data would not be identifiable in any report).
- ☐ The information you have submitted will be published as a report and you will be sent a copy upon request. Please note that confidentiality and anonymity will be maintained and it will not be possible to identify you from any publications.

Participant's Statement:

I _____ (full name, please print)

agree that the research project named above has been explained to me to my satisfaction and I agree to take part in the project. I have read both the notes written above and the Information Sheet about the project, and understand what the research involves.

Signed:

Date: _____

Appendix B: Participant Information Sheet

Information sheet for participants

Title of project: Finding My Voice: An Interdisciplinary and Multi-Methodological Investigation into the Relationship between Performers' Speech and Musical Expression

Study approved by School Research Ethics Committee:

I would like to invite you to participate in this PhD research project. You should only participate if you want to; choosing not to take part will not disadvantage you in any way. Before you decide whether you want to take part, it is important for you to understand why the research is being done and what your participation will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask me if there is anything that is not clear or if you would like more information.

- ☐ This research will try to find a connection between classical pianists' native language and their musical expression during performance. I will be observing pianists with Standard Mandarin Chinese and English native languages. I will analyse the recordings and speech of pianists in three different language groups: Chinese monolinguals (those who mainly speak Chinese in everyday life), English monolinguals, and Chinese/English bilinguals (fluent speech in both Chinese and English). I hope that with this experiment, I can provide evidence to support the idea that language plays a significant part in a musician's expression.
- ☐ The experiments will be conducted at the ** recording studio (B02). Members of ** audio/visual department will be present during the process, but they will all have agreed upon full discretion—no one will speak about the experiments outside of the experience.
- ☐ I am looking for classical pianists who fall into the following categories: 1. Standard Mandarin Chinese native speakers who mainly speak Chinese in everyday life, 2. English native speakers who speak only English in everyday life, 3. Chinese/English speakers who are fluent in both languages and are able to speak both in everyday life.
- ☐ If you would like to take part, you will be asked to give your availability. You will be given the excerpts to prepare as well as the spoken sentences to look over. The experiment itself will run a maximum of two hours per participant, including studio set up time and discussion of the instructions.
- ☐ You may be excluded from participation due to illness, emergency, and other unforeseen circumstances. It is possible to reschedule, but as I will have to book the recording studio, please let me know as soon as possible that you need to change the time.
- ☐ There are no risks involved in this experiment. Should you feel uncomfortable at any point during the process, you are able to stop and withdraw from participation.
- ☐ Possible benefits of this experiment include the creation of a data report relating speech to performance expression, which has previously never been made. Also, results may contribute to the literature of piano practice and be developed for educational purposes.
- ☐ Please remember: your data will be anonymised and you will not be identifiable by anyone besides myself. Your identity will be concealed in the report, as well as any form of raw data. Only myself and my supervisors as well as some audio/visual staff will have access to the data.
- ☐ If you wish to withdraw at any time, or decide not to take part, it will not affect the standard of education you receive at the **.

- ☐ If you decide to participate, you will be given this information sheet to keep and be asked to sign a consent form.
- ☐ Should you require or desire, I will be able to ask questions from the questionnaire in an interview format. Interviews will be recorded, subject to your permission. Recordings of interviews will be deleted upon transcription.
- ☐ The research outcomes will be available as a report, as well as incorporated into my final thesis. You will be given copies of the report and thesis upon request.

It is up to you to decide whether to take part or not. If you decide to take part you are still free to withdraw at any time and without giving a reason.

If this study has harmed you in any way you can contact the ** using the details below for further advice and information:

Appendix C: Questionnaires

Language Ability Questionnaire Part 1

PLEASE NOTE- you can stop at any time during this survey should you feel uncomfortable with the content!

Fluency= Ease; smoothness or flow; when speaking quickly, sounds, syllables, words, and phrases are easy to produce

Participant code (I will give to you):

Initials (only for my reference):

Date:

Age:

Date of Birth:

Gender:

1. Please list all the languages that you know, starting with the language that you feel *most comfortable* using:

2. Please list all the languages that you know, starting with the language you *learned first*:

3. Please list what percentage of time (on average) you are *currently exposed* to each language (*your percentages should add up to 100%*):

4. If you had to read a foreign text translated into all the languages that you know, in what percentage would you choose to read it in *each* of your languages (*your percentages should add up to 100%*)?

5. When choosing a language to speak with a person who is equally fluent in all your languages, what percentage of time would you choose to speak in each language (*your percentages should add up to 100%*)?

Identification scale (you can choose any number from 0-10):

0= this culture is completely unknown to me

1= I have a little knowledge of this culture

5= I have some knowledge of this culture and identify myself with some things from this culture

8= this culture feels familiar to me and I identify fairly strongly with it

10= this is the only culture I feel very familiar with, and I identify very strongly with it

6. Please name the cultures with which you identify (that you know and feel you belong in).

Example of cultures include American, British, Chinese, etc.:

7. Please rate (see identification scale) how much you identify yourself with each of your cultures:

8. How many years of formal education do you have? Please state your highest education level (High School Diploma, Bachelors, Masters, or PhD):

9. Date of arrival in the UK:

-if you have ever immigrated to another country, please provide the name of country and date:

10. Have you ever been diagnosed with a vision problem, hearing problem, language disability, or learning disability? (Check all that applies) If Yes, please explain (including corrections):

Language Ability Questionnaire Part 2 English

This is my _____ language. (options: first, second, third, etc.)

All questions below refer to your knowledge of _____ (state language).

1. What age did you begin *acquiring* the language?
 2. What age did you become *fluent* in the language?
 3. What age did you begin *reading* in the language?
 4. What age did you become *fluent in reading* the language?
5. Please list the number of years and/or months that you spent in each language environment:

-A country where the language is spoken:

-A family where the language is spoken:

-A school/working environment where the language is spoken:

On the following scale of **zero to ten**:

0: none

5: fairly fluent and somewhat comfortable

10: extremely fluent and comfortable

6. Please select the number of your level of proficiency in

-*speaking* the language:

-*understanding* the language:

-*reading* the language:

On the following scale:

0: none

5: fairly influential

10: extremely influential

7. Please select how much the following factors helped you to learn the language:

-spending time/talking with friends:

-spending time/talking with family:

-reading:

-language tapes/self-learning:

-watching TV:

-listening to radio/music:

-other:

On the following scale:

0: none

5: moderate amount

10: most significant amount

8. Please rate how much you are *currently* exposed to the language in the following ways:

-spending time/talking with friends:

-spending time/talking to family:

-reading:

-language tapes/self-learning:

-watching TV:

-listening to radio/music:

-other:

Language Ability Questionnaire Part 2 Mandarin Chinese

This is my _____ language. (options: first, second, third, etc.)

All questions below refer to your knowledge of _____ (state language).

1. What age did you begin *acquiring* the language?
 2. What age did you become *fluent* in the language?
 3. What age did you begin *reading* in the language?
 4. What age did you become *fluent in reading* the language?
5. Please list the number of years and/or months that you spent in each language environment:

-A country where the language is spoken:

-A family where the language is spoken:

-A school/working environment where the language is spoken:

On the following scale of **zero to ten**:

0: none

5: fairly fluent and somewhat comfortable

10: extremely fluent and comfortable

6. Please select the number of your level of proficiency in

-*speaking* the language:

-*understanding* the language:

-*reading* the language:

On the following scale:

0: none

5: fairly influential

10: extremely influential

7. Please select how much the following factors helped you to learn the language:

-spending time/talking with friends:

-spending time/talking with family:

-reading:

-language tapes/self-learning:

-watching TV:

-listening to radio/music:

-other:

On the following scale:

0: none

5: moderate amount

10: most significant amount

8. Please rate how much you are *currently* exposed to the language in the following ways:

-spending time/talking with friends:

-spending time/talking to family:

-reading:

-language tapes/self-learning:

-watching TV:

-listening to radio/music:

-other:

Questionnaire for Bilingual Background- English/Standard Mandarin Chinese

Participant Code:

Initials:

Age:

Gender:

I. Personal data-

1. What is the highest level of education completed?
2. Country of origin:
3. Country of current residence:
4. If you were not born in the U.K., during what ages did you live in your country of origin?
5. What other countries have you lived in and for how long?
6. If you were not born in the U.K., how long have you lived in the U.K. for?

II. Family history-

1. Where are your parents/caregivers from? Mother_____ Father_____
2. What languages do you parents/caregivers speak?
3. What do you parents do for a living?
4. What is your parents highest level of schooling?

III. Your linguistic history-

1. At what age did you first learn English?
2. At what age did you first learn Chinese?
3. Did you speak both Chinese and English before the age of 5? Yes, no
4. What languages did you *hear* in your home between the ages of birth-5 years?
Chinese, English, mixed, both, other (specify)

5. What languages did your parents/caregivers use mostly when speaking to you?
Chinese, English, mixed, both, other
6. What languages did you use mostly when speaking to your parents/caregivers?
Chinese, English, mixed, both, other
7. Do you have siblings? Yes, no, how many, are they older or younger?
8. What languages did you use to speak with your siblings?
9. What languages did your siblings use when speaking with you?
10. Did you play with other Chinese-speaking children?
11. What language did you use with other children/siblings?
12. Did you watch TV in Chinese?
13. Did your parents encourage you to speak as much Chinese as possible in the house?
14. Did your parents read/tell stories to you in Chinese?
15. Did your parents correct you when you spoke Chinese?

IV. Your linguistic proficiency NOW-

1. Rate your current language ability in ENGLISH by choosing a number:

- 1 = understand but cannot speak
- 2 = understand and can speak with great difficulty
- 3 = understand and speak but with some difficulty
- 4 = understand and speak comfortably, with little difficulty
- 5 = understand and speak fluently like a native speaker

2. Rate your current overall language ability in CHINESE:

- 1 = understand but cannot speak
- 2 = understand and can speak with great difficulty
- 3 = understand and speak but with some difficulty
- 4 = understand and speak comfortably, with little difficulty
- 5 = understand and speak fluently like a native speaker

3. On a scale from 1 to 5, rate your abilities in English and in Chinese
(1 =poor; 2= needs work; 3=good; 4= very good; 5= native speaker command)

English Reading =	Speaking=	Listening=	Writing=
Chinese	Reading =	Speaking=	Listening=
			Writing=

4. In general, as a young adult, which language do you prefer to use? (circle one)

English

Chinese

It depends

Both

on whom I talk to

5. Do you feel Chinese is your native language or like a second language?

Native language

second language

Appendix D: Participant Profiles

Man 1

Age: **21**

Gender: **Female**

Date of participation: **January 2017**

Languages known: **Chinese, English**

Order of languages learnt: **Chinese, English**

First language: **Chinese**

Second language: **English**

Place of birth: **China**

Date of arrival to the UK: **2015**

Never immigrated to another country.

Spent **2 years** in an English-speaking country.

On the following scale of one to five:

1: none

3: fairly fluent and somewhat comfortable

5: extremely fluent and comfortable

Self-rated level of proficiency in understanding and speaking English: **3**

Man 2

Age: **20**

Gender: **Female**

Date of participation: **January 2017**

Languages known: Chinese, English (German)

Order of languages learnt: **Chinese, English (German)**

First language: **Chinese**

Second language: **English**

Place of birth: **China**

Date of arrival to the UK: **2012**

Never immigrated to another country.

Spent **5 years** in an English-speaking country.

Self-rated level of proficiency in understanding and speaking English: **5**

Man 3

Age: **23**

Gender: **Female**

Date of participation: **May 2017**

Languages known: Chinese, English

Order of languages learnt: **Chinese, English**

First language: **Chinese**

Second language: **English**

Place of birth: **China**

Date of arrival to the UK: **2004**

Moved back to China in 2006, moved back to UK in 2010

Spent **9 years** in an English-speaking country.

Self-rated level of proficiency in understanding and speaking English: **4**

Man 4

Age: **24**

Gender: **Male**

Date of participation: **May 2018**

Languages known: Chinese, English

Order of languages learnt: **Chinese, English**

First language: **Chinese**

Second language: **English**

Place of birth: **China**

Date of arrival to the UK: **2017**

Date of arrival to the US: **2011**

Spent **6 years** in an English-speaking country.

Self-rated level of proficiency in understanding and speaking English: **4**

Man 5

Age: **21**

Gender: **Female**

Date of participation: **May 2018**

Languages known: Chinese, English

Order of languages learnt: **Chinese, English**

First language: **Chinese**

Second language: **English**

Place of birth: **China**

Date of arrival to the UK: **2016**

Would come to UK once a year for holiday prior

Spent **2 years** in an English-speaking country.

Self-rated level of proficiency in understanding and speaking English: **3**

Eng 1

Age: **27**

Gender: **Male**

Date of participation: **February 2017**

Languages known: **English**

First language: **English**

Place of birth: **UK**

Never immigrate to another country

Identifies as: **British**

Eng 2

Age: **25**

Gender: **Male**

Date of participation: **May 2017**

Languages known: English (Spanish)

First language: **English**

Place of birth: **US**

Date of arrival to the UK: **2016**

Never immigrated to another country

Identifies as: **American and Jewish**

Eng 3

Age: **23**

Gender: **Male**

Date of participation: **May 2017**

Languages known: **English**

First language: **English**

Place of birth: **US**

Date of arrival to the UK: **2016**

Never immigrated to another country

Identifies as: **American**

Eng 4

Age: **32**

Gender: **Female**

Date of participation: **June 2018**

Languages known: **English**

First language: **English**

Place of birth: **US**

Date of arrival to the UK: **2016**

Never immigrated to another country

Identifies as: **American**

Eng 5

Age: **21**

Gender: **Male**

Date of participation: **June 2018**

Languages known: English (Chinese)

First language: English (and Chinese until age 5)

Dominant language: **English**

Place of birth: **UK**

Never immigrated to another country

Identifies as: **British, with a little knowledge of Chinese culture**

Eng 6

Age: **25**

Gender: **Male**

Date of participation: **July 2018**

Languages known: **English**

First language: **English**

Place of birth: **Belgium**

Identifies as: **British, Belgian**

Appendix E: Mandarin Chinese Simple Story

chuāng wài

xiǎode shíhou, tīng guò yī gè gùshi。

gùshi zhōng yǒu liǎng gè niánqīng nǚzǐ。

yǒu yī gè yǒu xīnzàngbìng, jiào xiǎo méi。

lìngyī gè hěn jiànkāng, jiào zhēn zhēn。

yǒu yī tiān, xiǎo méi shuāi le yī jiǎo, yòu tuǐ duàn le, zhù jìn le yīyuàn。

ér zhēn zhēn yě zài tóngyī tiān chū chēhuò, liǎng tiáo tuǐ dōu duàn le。

窗外

小的时候，听过一个故事。

故事中有两个年轻女子。

有一个有心脏病，叫小梅。

另一个很健康，叫珍珍。

有一天，小梅摔了一跤，右腿断了，住进了医院。

而珍珍也在同一天出车祸，两条腿都断了。

Outside of the Window (English Translation)

When I was little, I heard about a story.

There were two young ladies in the story.

One of them was called Xiao Mei, who had heart problems.

The other one was very healthy, called Zhen Zhen.

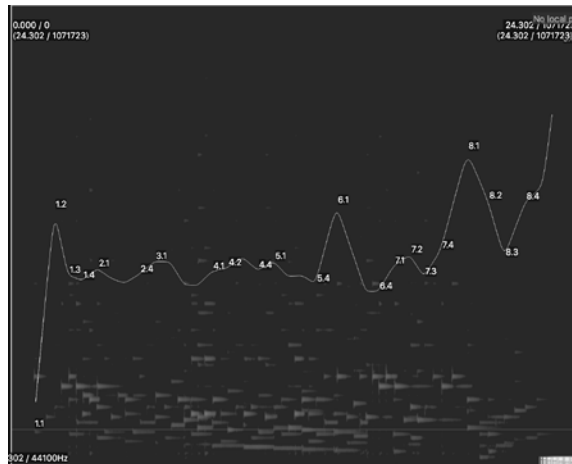
One day, Xiao Mei had a fall and broke her right leg. She was sent to the hospital after that.

Zhen Zhen was injured in a car accident on the same day, had both of her legs broken.

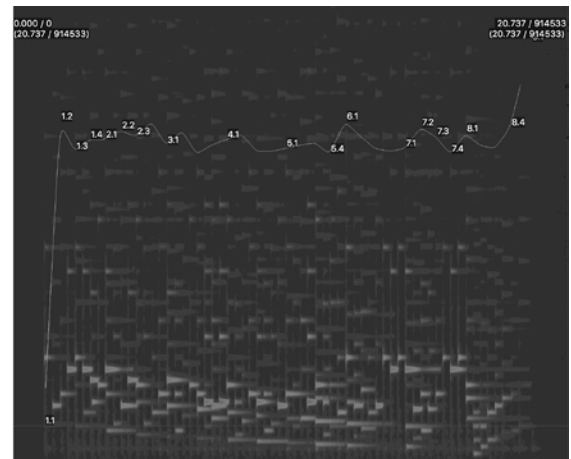
Appendix F: Timing Curves

Man 1

Expressive

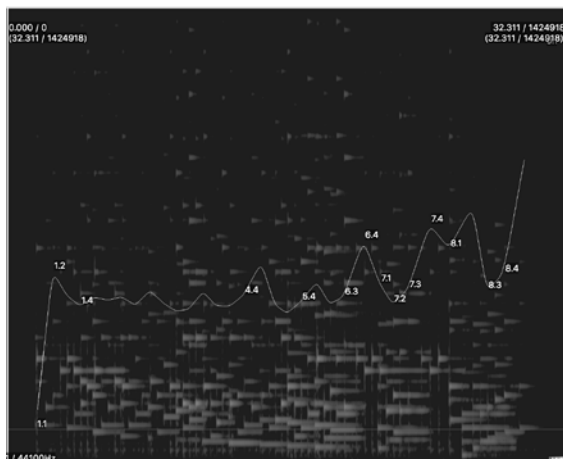


Mechanical

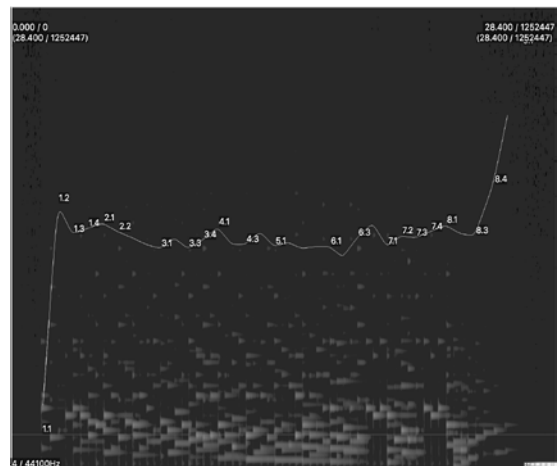


Man 2

Expressive

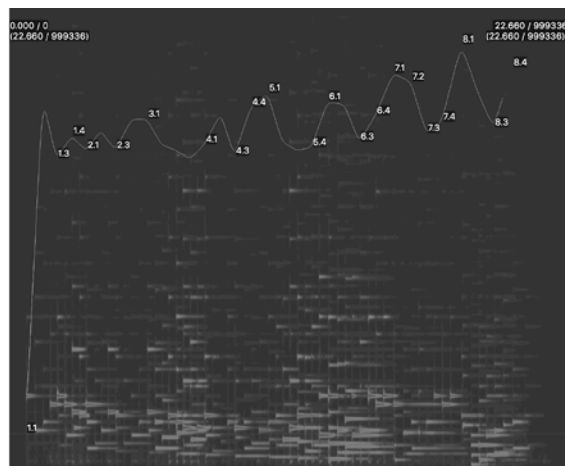


Mechanical

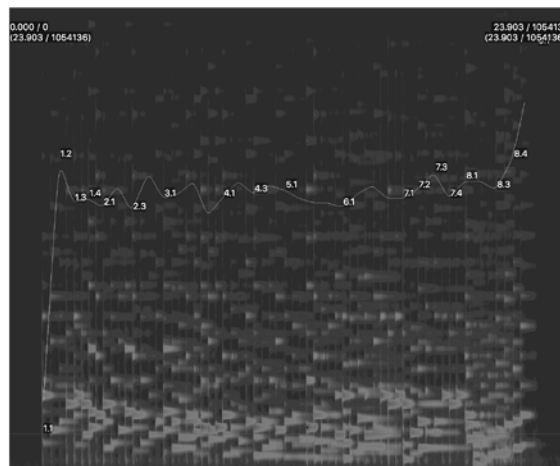


Man 3

Expressive

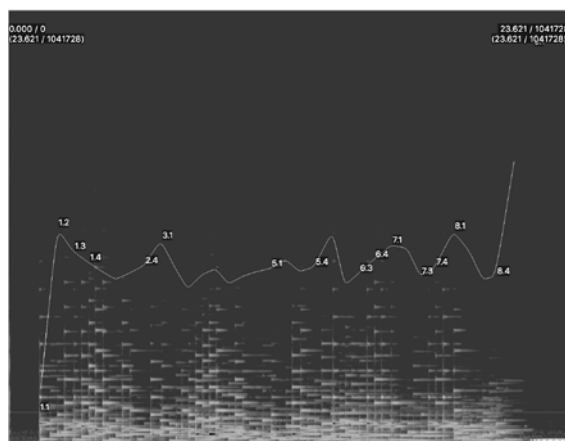


Mechanical

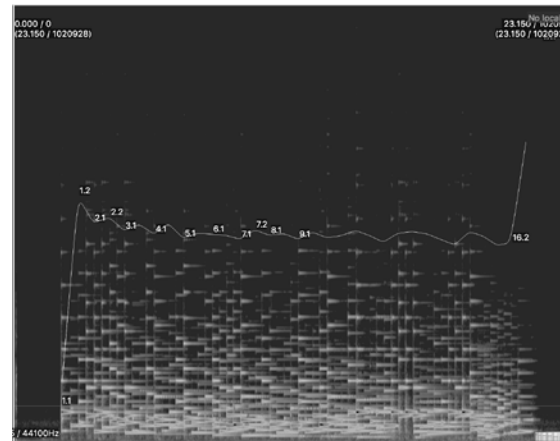


Man 4

Expressive

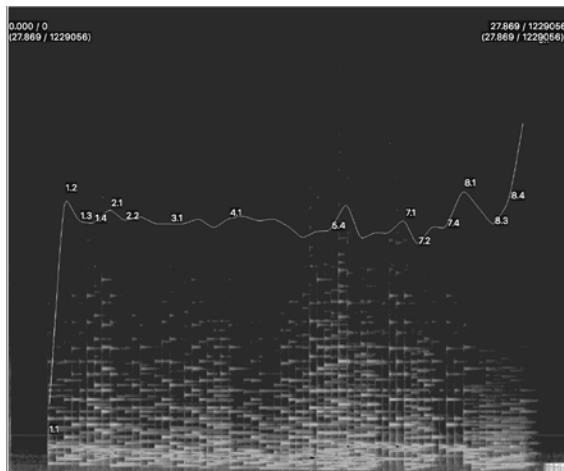


Mechanical

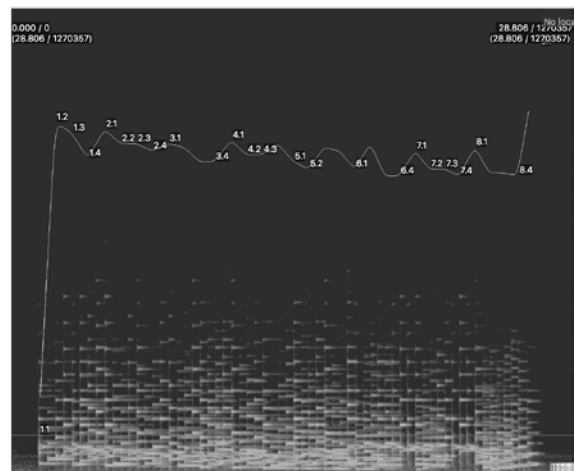


Man 5

Expressive

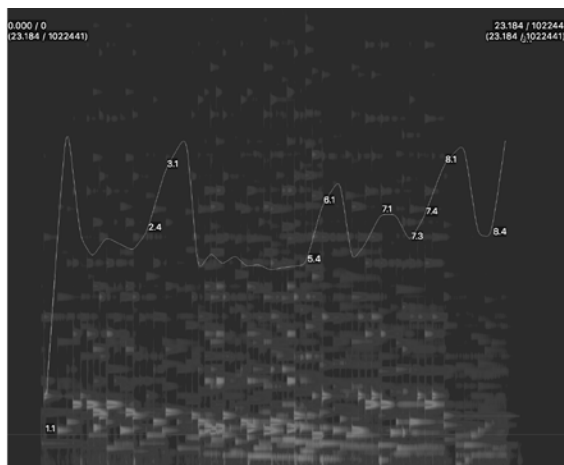


Mechanical

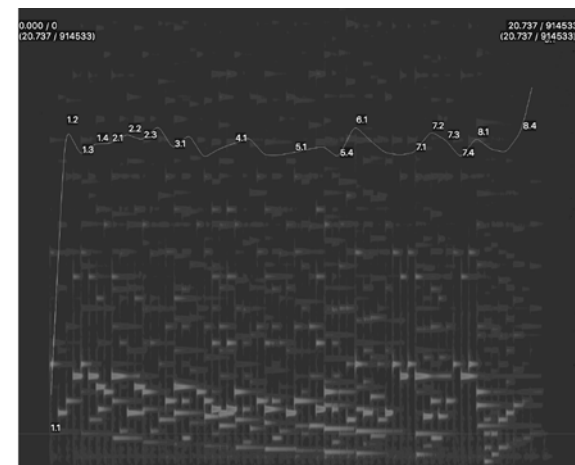


Eng 1

Expressive

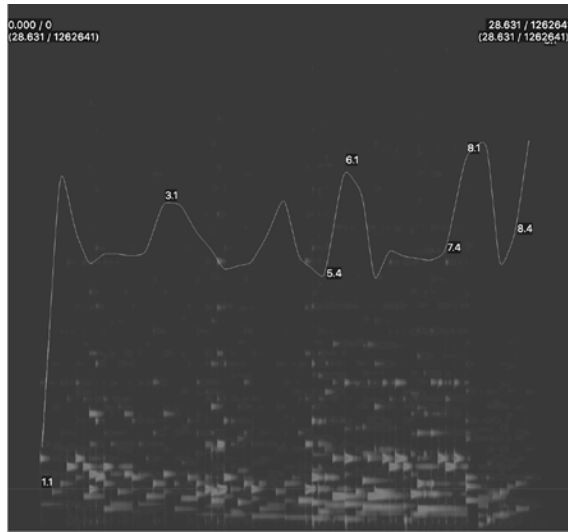


Mechanical

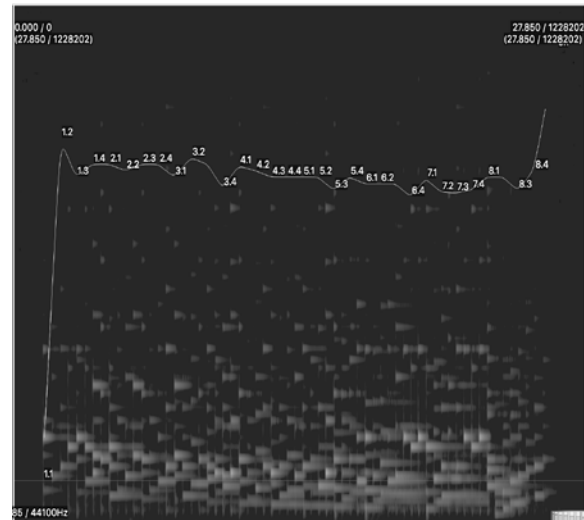


Eng 2

Expressive



Mechanical

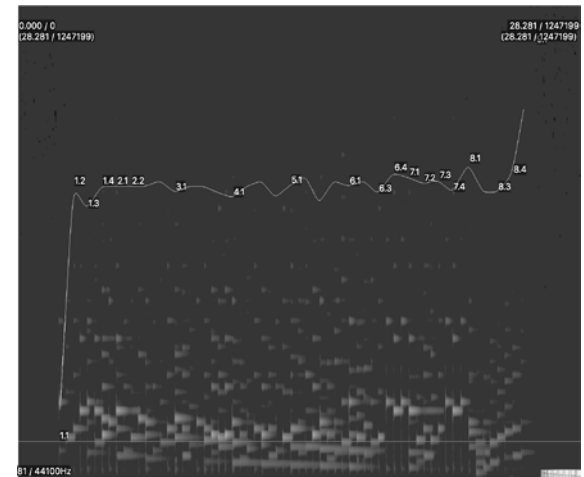


Eng 3

Expressive

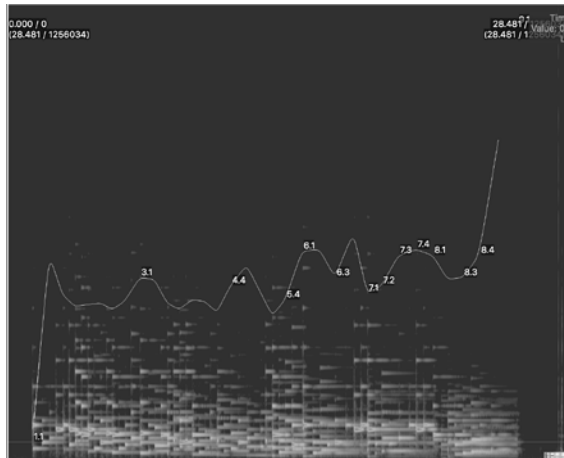


Mechanical

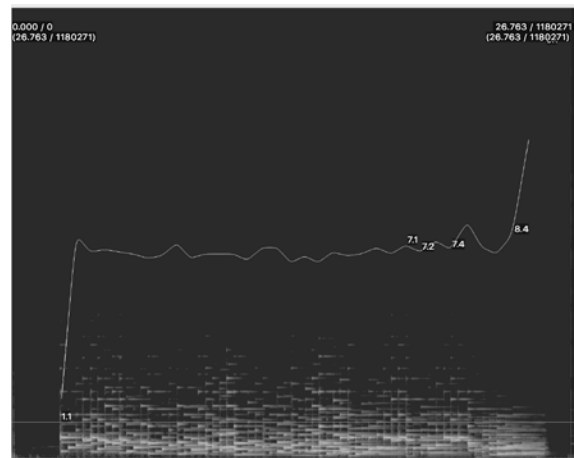


Eng 4

Expressive

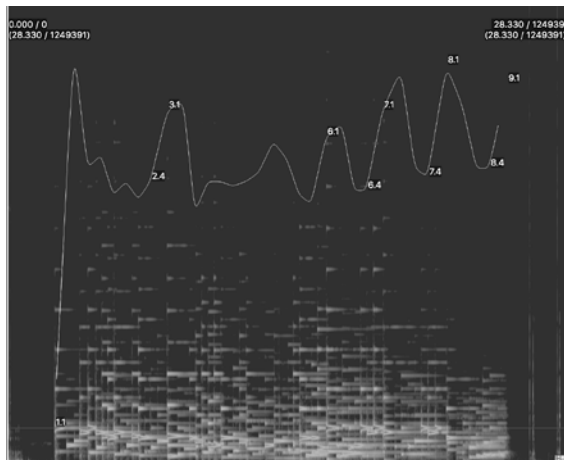


Mechanical

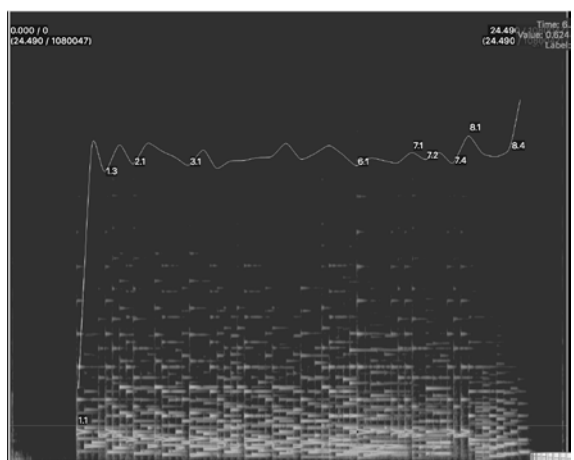


Eng 5

Expressive

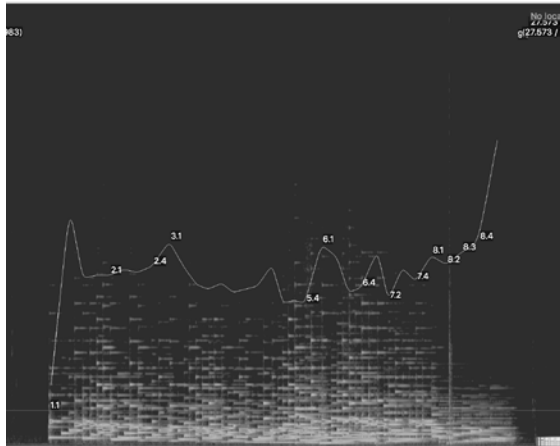


Mechanical

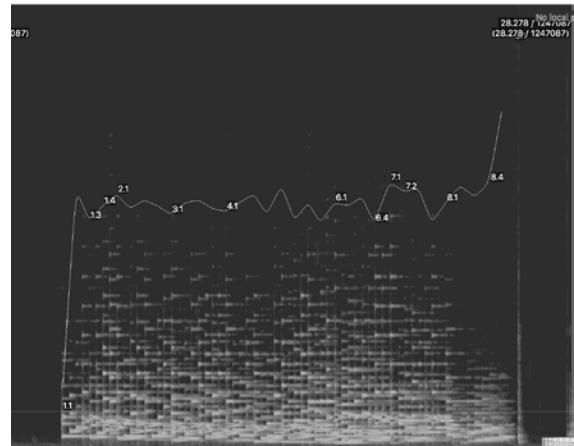


Eng 6

Expressive



Mechanical



Bibliography

- Abercrombie, David. 1967. *Elements of General Phonetics* (Edinburgh: University Press)
- Abraham, Gerald. 1974. *The Tradition of Western Music* (London: Oxford University Press)
- Adorno, Theodor W. *Alban Berg, Master of the Smallest Link*, trans. by Juliane Brand and Christopher Hailey (Cambridge: Cambridge University Press)
- Akbar, Shawn R. 2012. 'Musical Understanding: Studies in Philosophy and Phenomenological Psychology', (PhD thesis, University of Iowa)
- Akiyama, Tatsuhide. 1976. *Nihon No Yōgaku Hyakunen-Shi* (Tokyo: Dao-Ichi Hōki Shuppan)
- Allen, George D., and Sarah Hawkins. 1980. 'Phonological rhythm: Definition and development', in *Child Phonology: Volume 1. Production*, ed. by G. Yeni-Komshian, J. Kavanagh, and C. Ferguson (New York: Academic Press) 227–256
- Anderson, Leon. 2006. 'Analytic Autoethnography', *Journal of Contemporary Ethnography*, 35, 373–395
- Anvari, Sima H, Laurel J Trainor, Jennifer Woodside, and Betty Ann Levy. 2002. 'Relations Among Musical Skills, Phonological Processing, And Early Reading Ability in Preschool Children', *Journal of Experimental Child Psychology*, 83, 111–130
- Arbib, Michael. 2002. 'The Mirror System, Imitation, and the Evolution of Language', in *Mitation In Animals and Artifacts* (Cambridge, MA: MIT Press) 229–280
- Arbib, Michael. 2005. 'An Action-Oriented Neurolinguistic Framework for the Evolution of Protolanguage', in *Language Origins: Perspectives on Evolution* (Oxford: Oxford University Press) 21–47
- Arvaniti, Amalia. 2009. 'Rhythm, Timing and the Timing of Rhythm', *Phonetica*, 66, 46–63

- Arvaniti, Amalia. 2012. 'The Usefulness of Metrics in the Quantification of Speech Rhythm', *Journal of Phonetics*, 40, 351–373
- Audacity Team. 2019. *Audacity(R): Free Audio Editor and Recorder*, version 2.1.1 (<http://audacity.sourceforge.net/>)
- Baker, Nancy K., Max Paddinson, and Roger Scruton. 2001. 'Expression' in *Grove Music Online* (Oxford Music Online)
<<https://www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000009138>> [Accessed 7 August 2019]
- Barlow, Harold, and Sam Morgenstern. 1983. *A Dictionary of Musical Themes* (London: Faber and Faber)
- Bartleet, Brydie-Leigh and Carolyn Ellis. 2009. 'Introduction: Making Autoethnography Sing/ Making Music Personal', in *Music Autoethnographies*, ed. by Brydie-Leigh Bartleet and Carolyn Ellis (Bowen Hills, QLD: Australian Academic Press) 1–20
- Bartleet, Brydie-Leigh. 2009. 'Behind the Baton: Exploring Autoethnographic Writing in A Musical Context', *Journal of Contemporary Ethnography*, 38, 713–733
- Beckman, Mary E. 1992. 'Evidence for Speech Rhythms Across Languages', in *Speech perception, production, and linguistic structure*, ed. by Yoh'ichi Tohkura, Eric Vatikiotis-Bateson, and Yoshinori Sagisaka (Tokyo: Ohmsha) 458–463
- Beckman, Mary E. 1996. 'The Parsing of Prosody', *Language and Cognitive Processes*, 11, 17–68
- Benjamin, Walter. 1969. 'The Work of Art in the Age of Mechanical Reproduction', in *Illuminations: Essays and Reflections* (New York: Schocken books), 166–195
<<https://web.mit.edu/allanmc/www/benjamin.pdf>> [Accessed 28 July 2019]
- Benton, Matthew, Liz Dockendorf, Wenhua Jin, Yang Liu, and Jerry Edmondson. 2007. 'The Continuum of Speech Rhythm: Computational Testing of Speech Rhythm of Large

- Corpora from Natural Chinese and English Speech', in *Proceedings of the 16th International Congress of Phonetic Sciences* (Saarbrücken, Germany: Universität des Saarlandes), 1269-1272 <<http://www.icphs2007.de>>
- Bernstein, Leonard. 1976. *The Unanswered Question* (Cambridge, MA: Harvard University Press)
- Bever, Thomas G. and Robert J. Chiarello. 1974. 'Cerebral Dominance in Musicians and Nonmusicians', *Science*, 185, 537–539
- Boersma, Paul, and David Weenink. 2019. 'Praat: Doing Phonetics by Computer' <<http://www.fon.hum.uva.nl/praat/>> [Accessed 6 August 2018]
- Bolinger, Dwight. 1981. *Two Kinds of Vowels, Two Kinds of Rhythm* (Bloomington: Indiana University Linguistics Club)
- Bolton, Thaddeus L. 1894. 'Rhythm', *The American Journal of Psychology*, 6, 145
- Boswell, George W. 1977. 'Pitch: Musical and Verbal in Folksong', *Yearbook of the International Folk Music Council*, 9, 80–88
- Bradlow, Ann R., Lynne C. Nygaard, and David B. Pisoni. 1999. 'Effects of Talker, Rate, And Amplitude Variation on Recognition Memory for Spoken Words', *Perception & Psychophysics*, 61, 206–219
- Bridge, Frank. 1915. *3 Sketches, H. 68* (London: Winthrop Rogers)
- Brown, Steven. 2000. 'The 'Musilanguage' Model of Music Evolution', in *The Origins of Music* (Cambridge: MIT Press) 271–300
- Buck, Bryony, Jennifer MacRitchie, and Nicholas J. Bailey. 2013. 'The Interpretive Shaping of Embodied Musical Structure in Piano Performance', *Empirical Musicology Review*, 8, 92

- Caballero, Carlo. 2003. *Fauré and French Musical Aesthetics* (Cambridge: Cambridge University Press)
- Cannam, Chris, Christian Landone, and Mark Sandler. 2010. 'Sonic Visualiser: An Open Source Application for Viewing, Analysing, and Annotating Music Audio Files', in *ACM Multimedia 2010 International Conference* (New York: Proceedings of the ACM Multimedia 2010 International Conference), 1467–1468
- Carner, Mosco. 1983. *Alban Berg*, 2nd edn (New York: Holmes & Meier)
- Carpenter, Angela C., and Andrea G. Levitt. 2016. 'Rhythm in the Speech and Music of Jazz and Riddim Musicians', *Music Perception: An Interdisciplinary Journal*, 34, 94–103
- Chang, Heewon. 2008. 'Autoethnography', in *Autoethnography as method* (Walnut Creek: Left Coast Press) 43–57
- Chomsky, Noam. 2002. *Syntactic Structures*, 2nd edn (Berlin: Mouton de Gruyter)
- Chua, Amy. 2011. *Battle Hymn of the Tiger Mother* (New York: Penguin Publishing Group)
- Clarke, Eric F. 1989. 'Issues in Language and Music', *Contemporary Music Review*, 4, 9–22
- Clarke, Eric F. 1993. 'Imitating and Evaluating Real and Transformed Musical Performances', *Music Perception: An Interdisciplinary Journal*, 10, 317–341
- Clarke, Eric, and Nicholas Cook (eds.). 2004. *Empirical Musicology: Aims, Methods, Prospects* (New York: Oxford University Press)
- Clarke, Eric. 1999. 'Rhythm and Timing in Music' in *Academic Press series in cognition and perception: A series of monographs and treatises. The psychology of music*, ed. by Diana Deutsch (San Diego, CA: Academic Press) 473–500
- Clifford, James. 2002. *The Predicament of Culture* (Cambridge, MA: Harvard University Press)

- Clynes, Manfred. 1982. *Music, Mind, And Brain: The Neuropsychology of Music* (New York: Springer Science+Business Media)
- Cone, Edward T. 1974. *The Composer's Voice* (Berkeley, CA: University of California Press)
- Cook, Nicholas. 2001. 'Between Process and Product: Music and/as Performance', *Music Theory Online*, 7.2
<<http://www.mtosmt.org/issues/mto.01.7.2/mto.01.7.2.cook.html>> [Accessed 15 August 2019]
- Cook, Nicholas. 2003. 'Music as Performance' in *The Cultural Study of Music: A Critical Introduction*, ed. by Martin Clayton, Trevor Herbert, and Richard Middleton (New York: Routledge) 204–214
- Cook, Nicholas. 2007. 'Performance Analysis and Chopin's Mazurkas', *Musicae Scientiae*, 11, 183–207
- Cook, Nicholas. 2009. 'Methods for analysing recordings', in *The Oxford Companion to Recorded Music*, ed. by Nicholas Cook, Eric Clarke, Daniel Leech-Wilkinson, and John Rink (Cambridge: Cambridge University Press) 221–245
- Cook, Nicholas. 2013. *Beyond the Score: Music as Performance* (New York, NY: Oxford University Press)
- Cooke, Deryck. 1959. *Language of Music* (London: Oxford University Press)
- Cooper, Grosvenor W, and Leonard B Meyer. 1960. *The Rhythmic Structure of Music* (Chicago: University of Chicago Press)
- Couper-Kuhlen, Elizabeth. 1986. *An Introduction to English Prosody* (Tubingen: Niemeyer)
- Crafts, Susan D, Daniel Cavicchi, and Charles Keil. 1993. *My Music* (Hanover, NH: University Press of New England)

- Cutler, Anne. 1995. 'Spoken word recognition and production', in *Speech, language, and communication*, ed. by Joanne Miller and Peter Eimas (San Diego, CA: Academic Press) 97–136
- D., T. 2011. 'The Model Minority: Asian-American Youth and the Harmful Perpetuation of a Cultural Myth', *Inquiries Journal/Student Pulse*, 3
<<http://www.inquiriesjournal.com/articles/571/the-model-minority-asian-american-youth-and-the-harmful-perpetuation-of-a-cultural-myth>> [Accessed 28 July 2019]
- Darwin, Charles. 1871. *The Descent of Man, and Selection in Relation to Sex* (New York: D. Appleton and Company)
- Dauer, Rebecca M. 1983. 'Stress-timing and Syllable-timing Reanalysed', *Journal of Phonetics*, 11, 51–62
- Dauer, Rebecca M. 1987. 'Phonetic and Phonological Components of Language Rhythm', in *Proceedings of the 11Th International Congress of Phonetic Sciences* (Tallinn, Estonia), 447–450
- Davidson, Jane W. 2014. 'Implications for Empirical Expressive Music Performance Research', in *Expressiveness in music performance*, ed. by Dorottya Fabian, Renee Timmers, and Emery Schubert (Oxford: Oxford University Press) 344–347
- Davidson, Jane W. and Nicole Jordan. 2007. "'Private teaching, private learning?": An exploration of music instrument learning in the private studio, junior and senior conservatories', in *International Handbook of Research in Arts Education*, ed. by Liora Bresler (Dordrecht, The Netherlands: Springer) 729–744
- Davies, Stephen and Stanley Sadie. 2001. 'Interpretation' in *Grove Music Online* (Oxford Music Online)
<<https://www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000013863>> [Accessed 7 August 2019]

- Davies, Stephen. 1994. *Musical Meaning and Expression* (Ithaca: Cornell University Press)
- Davies, Stephen. 2001. 'IV. Anglo-American philosophy of music, 1960-2000', in 'Philosophy of Music', *Grove Music Online* (Oxford Music Online)
<<https://www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000052965>> [Accessed 7 August 2019]
- Day, Anthony. 1994. 'A Shift in Composition', *The Los Angeles Times*
- Dellwo, Volker, Bianca Aschenberger, Petra Wagner, Jana Dancovicova, and Ingmar Steiner. 2004. 'Bonntempo-Corpus and Bonntempo-Tools: A Database for the Study of Speech Rhythm and Rate', in *8th International Conference on Spoken Language Processing* (Jeju Island, Korea: INTERSPEECH-2004), 777–780
- Dellwo, Volker. 2006. 'Rhythm and Speech Rate: A Variation Coefficient Of ΔC ', in *Language and Language-Processing: Proceedings of the 38th Linguistics Colloquium, Piliscsaba 2003*, ed. by Pawel Karnowski and Imre Syigeti (Frankfurt am Main, Germany: Peter Lang), 231–241
- DeNora, Tia. 2013. *Music in Everyday Life* (Cambridge: Cambridge University Press)
- Denshire, Sally. 2013. 'Autoethnography', *Sociopedia*
<<http://www.sagepub.net/isa/resources/pdf/Autoethnography.pdf>> [Accessed 14 August 2019]
- Deutsch, Diana, Tom North, and Lee Ray. 1990. 'The Tritone Paradox: Correlate with the Listener's Vocal Range for Speech', *Music Perception: An Interdisciplinary Journal*, 7, 371–384
- Deutsch, Diana. 1990. 'A Link Between Music Perception and Speech Production', *Journal of the Acoustical Society of America*, 88, S139–S139
- Deutsch, Diana. 1991. 'The Tritone Paradox: An Influence of Language on Music Perception', *Music Perception*, 8, 335–347

- Deutsch, Diana. 1999. 'The Processing of Pitch Combinations', in *The Psychology of Music*, 2nd edn (Academic Press) 349–411
- Dobszay, L. 1972. 'The Kodály Method and Its Musical Basis', *Studia Musicologica Academiae Scientiarum Hungaricae* 14, 15–33
- Doğantan-Dack, Mine. 2008. 'Recording the performer's voice', in *Recorded Music: Philosophical and Critical Reflections*, ed. by Mine Doğantan-Dack (London: Middlesex University Press) 293–313
- Doğantan-Dack, Mine. 2012a. 'The Art of Research in Live Music Performance', *Music Performance Research, Special Issue*, 5, 34–48
- Doğantan-Dack, Mine. 2012b. "'Phrasing – The Very Life of Music": Performing the Music and Nineteenth-Century Performance Theory', *Nineteenth-Century Music Review*, 9, 7–30
- Doğantan-Dack, Mine. 2015. 'Artistic Research in Classical Music Performance: Truth and Politics', *PARSE – Journal of Art and Research*, Spring
<<http://parsejournal.com/article/artistic-research-in-classical-music-performance/>>
[Accessed 15 August 2019]
- Dowling, W. Jay, and Dane L Harwood. 1986. *Music Cognition* (New York: Academic Press)
- Duckles, Vincent and Jann Pasler. 2001. 'I. The nature of musicology', in 'Musicology', *Grove Music Online* (Oxford Music Online)
<<https://www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000046710>> [Accessed 9 August 2019]
- Eigeldinger, Jean-Jacques. 1986. *Chopin – Pianist and Teacher, as seen by his pupils*, trans. by Naomi Sholet with Krysia Osostowicz and Roy Howat, ed. by Roy Howat (Cambridge: Cambridge University Press)

- Ellis, Carolyn and Arthur P. Bochner. 2000. 'Autoethnography, Personal, Narrative, Reflexivity: Researcher as subject', in *Handbook of Qualitative Research, 2nd Edition*, ed. by Norman Denzin and Yvonna Lincoln (Thousand Oaks, CA: Sage Publications) 733–768
- Ellis, Carolyn. 1991. 'Sociological introspection and emotional experience', *Symbolic Interaction*, 14(1), 23–50
- Ellis, Carolyn. 1999. 'Heartful Autoethnography', *Qualitative Health Research*, 9, 669–683
- Emirbayer, Mustafa, and Ann Mische. 1998. 'What is Agency?', *American Journal of Sociology*, 103, 962–1023
- Emmerson, Stephen. 2009. 'Evoking spring in winter: Some personal reflections on returning to Schubert's cycle', in *Music Autoethnographies*, ed. by Brydie-Leigh Bartleet and Carolyn Ellis (Bowen Hills, QLD: Australian Academic Press) 101–120
- Eppstein, Ury. 1994. *The Beginnings of Western Music in Meiji Era Japan* (Lewiston, N.Y: Edwin Mellen Press)
- Evans, James R., and Manfred Clynes. 1986. *Rhythm in Psychological, Linguistic and Musical Processes* (Springfield, IL: Charles C Thomas)
- Fabian, Dorottya, Renee Timmers, and Emery Schubert. 2014. *Expressiveness in Music Performance: Empirical Approaches Across Styles and Cultures* (Oxford: Oxford University Press)
- Feld, Steven, and Aaron A. Fox. 1994. 'Music and Language', *Annual Review of Anthropology*, 23, 25–53
- Feld, Steven. 1989. 'Sound', in *International Encyclopedia of Communications* (Oxford: Oxford University Press) 101–107

- Ferreira, Fernanda. 2002. 'Prosody', in *Encyclopedia of Cognitive Science*, ed. by Lynn Nadel (New York: Nature Publishing) 258–265
- Ford, Biranda. 2010. 'What Are Conservatoires For? Discourses of Purpose in The Contemporary Conservatoire' (PhD thesis, Institute of Education, University of London)
- Gabriel, Clive. 1978. 'An Experimental Study of Deryck Cooke's Theory of Music and Meaning', *Psychology of Music*, 6, 13–20
- Gabrielsson, Alf, and Patrik N. Juslin. 1996. 'Emotional Expression in Music Performance: Between the Performer's Intention and the Listener's Experience', *Psychology of Music*, 24, 68–91
- Gabrielsson, Alf. 1973. 'Adjective Ratings and Dimension Analyses of Auditory Rhythm Patterns', *Scandinavian Journal of Psychology*, 14, 244–260
- Gabrielsson, Alf. 1987. 'Once again: The theme from Mozart's Piano Sonata in A major: A comparison of five performances', in *Action and Perception in Rhythm and Music*, ed. by Alf Gabrielsson (Stockholm: Royal Swedish Academy of Music) 81–103
- Garfias, Robert. 1987. 'Thoughts on the Process of Language and Music Acquisition', in *Music and Child Development: Proceedings of the 1987 Biology of Music Making Conference*, ed. by Franz L. Roehmann and Frank R. Wilson (St. Louis, MO: MMB Music) 100–105
- Gerbino Giuseppe and Iain Fenlon. 2006. 'Early Opera: The Initial Phase', in *European Music, 1520-1640*, ed. by James Haar (Woodbridge: The Boydell Press) 472–486
- Gertsch, Emily S. 2013. 'Narratives of Innocence and Experience: Plot Archetypes in Robert Schumann's Piano Quintet and Piano Quartet' (PhD thesis, Florida State University)
- Giuliano, Ryan J., Peter Q. Pfordresher, Emily M. Stanley, Shalini Narayana, and Nicole Y. Y. Wicha. 2011. 'Native Experience with A Tone Language Enhances Pitch

Discrimination and the Timing of Neural Responses to Pitch Change', *Frontiers in Psychology*, 2, 146

Godlovitch, Stanley. 1998. *Musical Performance: A Philosophical Study* (London: Routledge)

Goehr, Lydia, F. E. Sparshott, Andrew Bowie, and Stephen Davies. 2001. 'Philosophy of music', in *Grove Music Online* (Oxford Music Online)
<<https://www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000052965>> [Accessed 7 August 2019]

Goehr, Lydia. 1992. *The Imaginary Museum of Musical Works: An Essay in the Philosophy of Music* (New York: Oxford University Press)

Goodman, Nelson. 1968. *Languages of Art: An Approach to A Theory of Symbols* (Indianapolis, IN: Bobbs-Merrill Company)

Gordon, Reyna L., Cyrille L. Magne, and Edward W. Large. 2011. 'EEG Correlates of Song Prosody: A New Look at the Relationship Between Linguistic and Musical Rhythm', *Frontiers in Psychology*, 2

Gussenhoven, Carlos. 2001. 'Intonation and biology', in *Liber Amicorum Bernard Bichakjian*, ed. by Haïke Jacobs and W. Leo Wetzels (Maastricht: Shaker) 59–82

Guter, Eran. 2004. 'Where Languages End: Ludwig Wittgenstein at the Crossroads of Music, Language, and the World' (PhD thesis, Boston University)

Hall, R. A. 1953. 'Elgar and the intonation of British English', *Gramophone*, 31, 6, reprinted in *Intonation: selected readings*, ed. by D. Bolinger (Hammondsworth: Penguin, 1972) 282–285

Handel, Stephen. 1989. *Listening: An Introduction to the Perception of Auditory Events* (Cambridge: A Bradford Book)

Harich-Schneider, Eta. 1973. *A History of Japanese Music* (London: Oxford University Press)

- Hawkins, Sarah, Ian Cross, and Richard Ogden. 2013. 'Communicative interaction in spontaneous music and speech', in *Language, Music and Interaction*, ed. by Martin Orwin, Christine Howes, and Ruth Kempson (London: College Publications) 285–329
- Hawkins, Sarah. 2014. 'Situational Influences on Rhythmicity in Speech, Music, And Their Interaction', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369, 20130398
- Hayano, David. 1979. 'Auto-Ethnography: Paradigms, Problems, And Prospects', *Human Organization*, 38, 99–104
- Heffner, Roe-Merrill Secrist. 1975. *General Phonetics* (Madison: Wisconsin University Press)
- Hegel, Georg W. F. 1975. *Aesthetics: Lectures on Fine Art*, trans. by Thomas M. Knox (New York: Oxford University Press)
- Henderson, Mack T. 1936. 'Rhythmic organization in artistic piano performance' in *Objective analysis of musical performance, Vol. 4*, ed. by Carl Seashore (Iowa City: University of Iowa Press) 281–305
- Herndon, Marcia. 1974. 'Analysis: The Herding of Sacred Cows?', *Ethnomusicology*, 18, 219–62
- Hilton, Leanne. 1990. 'Song Metrics', in *The Annual Proceedings of the Berkeley Linguistics Society: Special Session on General Topics in American Indian Linguistics* (Berkeley, CA: University of California, Berkeley) 51–60
<<https://journals.linguisticsociety.org/proceedings/index.php/BLS/article/viewFile/1678/1452>>
- Ho, Yim-Chi, Mei-Chun Cheung, and Agnes S. Chan. 2003. 'Music Training Improves Verbal but Not Visual Memory: Cross-Sectional and Longitudinal Explorations in Children.', *Neuropsychology*, 17, 439–450

- Hobsbawm, Eric, and Terence Ranger (ed.). 1983. *The Invention of Tradition* (Cambridge: Cambridge University Press)
- Hoch, Lisianne, Benedicte Poulin-Charronnat, and Barbara Tillmann. 2011. 'The Influence of Task-Irrelevant Music on Language Processing: Syntactic and Semantic Structures', *Frontiers in Psychology*, 2, 112
- Hoffmann, E.T.A. 1989. 'Kreisleriana: Extremely Random Thoughts', in *E.T.A. Hoffmann's Musical Writings*, ed. by David Charlton (Cambridge: Cambridge University Press)
- Houghton, Catherine. 1984. 'Structure in Language and Music: A Linguistic Approach' (PhD thesis, Stanford University)
- Hullah, Annette. 2009. *Theodor Leschetizky* (Whitefish, MT: Kessinger Publishing)
- Huron, David, and Joy Ollen. 2003. 'Agogic Contrast in French and English Themes: Further Support for Patel And Daniele (2003)', *Music Perception*, 21, 267–271
- Hwang, Okon. 2001. 'Western Art Music in Korea: Everyday Experience and Cultural Critique' (PhD thesis, Wesleyan University)
- Imberty, Michel. 1975. 'Perspectives Nouvelles De La Sémantique Musicale Expérimentale', *Musique En Jeu*, 17, 87–109
- Iversen, John R. 2008. 'Review Of 'Perception and Production of Linguistic and Musical Rhythm by Korean And English Middle School Students', by Lydia N. Slobodian', *Empirical Musicology Review*, 3, 208–214
- Iversen, John R., Aniruddh D. Patel, and Kengo Ohgushi. 2008. 'Perception of Rhythmic Grouping Depends on Auditory Experience', *The Journal of the Acoustical Society of America*, 124, 2263–2271
- Jackendoff, Ray. 1977. review of Leonard Bernstein, *The Unanswered Question* (1976), *Language*, 53, 883–894

- Jackendoff, Ray. 1989. 'A Comparison of Rhythmic Structures in Music and Language', in *Phonetics and Phonology: Rhythm and Meter*, 1st edn (San Diego, CA: Academic Press) 15–44
- Jackendoff, Ray. 2009. 'Parallels and Nonparallels Between Language and Music', *Music Perception: An Interdisciplinary Journal*, 26, 195–204
- Jakobson, Lorna S., Samantha T. Lewycky, Andrea R. Kilgour, and Brenda M. Stoesz. 2008. 'Memory for Verbal and Visual Material in Highly Trained Musicians', *Music Perception*, 26, 41–55
- Jakobson, Roman. 1987. *Language in Literature*, ed. by Krystyna Pomorska and Stephen Rudy (Cambridge: Harvard University Press)
- Jentschke, Sebastian. 2016. 'The Relationship between Music and Language', in *The Oxford Handbook of Music Psychology*, 2nd edition, ed. by Susan Hallam, Ian Cross, and Michael Thaut (Oxford: Oxford University Press)
- Jeon, Hae-Sung, and Francis Nolan. 2013. 'The Role of Pitch and Timing Cues in the Perception of Phrasal Grouping in Seoul Korean', *The Journal of the Acoustical Society of America*, 133, 3039–3049
- Jeon, Hae-Sung. 2011. 'Prosodic Phrasing in Seoul Korean: The Role of Pitch and Timing Cues' (PhD thesis, University of Cambridge)
- Johnson, Ian. 2010. 'Going Baroque in China', *The New York Times*, <<https://www.nytimes.com/2010/10/28/arts/28iht-baroque.html>> [Accessed 13 August 2019]
- Jones, Andrew F. 2001. *Yellow Music: Media Culture and Colonial Modernity in the Chinese Jazz Age* (Durham: Duke University Press)

- Juslin, Patrik N. 1998. 'A functionalist perspective on emotional communication in music performance' (PhD thesis, Uppsala University) in *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences*, 78, 7–65
- Juslin, Patrik N., and Petri Laukka. 2003. 'Communication of Emotions in Vocal Expression and Music Performance: Different Channels, Same Code?', *Psychological Bulletin*, 129, 770–814
- Kania, Andrew. 2013. 'Music', in *The Routledge Companion to Aesthetics*, third edition, ed. by Berys Gaut and Dominic McIver Lopes (New York: Routledge) 639–48
- Keating, Patricia, and Grace Kuo. 2012. 'Comparison of Speaking Fundamental Frequency in English And Mandarin', *Journal of the Acoustical Society of America*, 132, 1050–1060
- Kendall, Roger A., and Edward C. Carterette. 1990. 'The Communication of Musical Expression', *Music Perception: An Interdisciplinary Journal*, 8, 129–163
- Kidd, Teri Duke. 2008. 'The Influence of the Hungarian Language and Hungarian Folk Song on a Performance of Selections from Fifteen Hungarian Peasant Songs by Bela Bartok' (PhD thesis, Longwood University)
- Kim, Kyung-Hee. 2016. *The Creativity Challenge: How We Can Recapture American Innovation* (Amherst, NJ: Prometheus Books)
- Kingsbury, Henry. 1988. *Music, Talent, And Performance: A Conservatory Cultural System* (Philadelphia: Temple University Press)
- Kiparsky, Peter, and Gilbert Youmans. 1989. *Phonetics and Phonology: Rhythm and Meter*, 1st edn (San Diego, CA: Academic Press)
- Kivy, Peter. 1980. *The Corded Shell: Reflections on Musical Expression* (Princeton, NJ: Princeton University Press)
- Kivy, Peter. 1989. *Sound Sentiment* (Philadelphia: Temple University Press)

- Kivy, Peter. 1990. *Music Alone: Philosophical Reflections on the Purely Musical Experience* (Ithaca, NY: Cornell University Press)
- Kivy, Peter. 1995. *Authenticities: Philosophical Reflections on Musical Performance* (Ithaca: Cornell University Press)
- Koelsch, Stefan. 2012. *Brain and Music* (Hoboken: John Wiley & Sons)
- Koelsch, Stefan. 2013. 'Neural Correlates of Music Perception', in *Language, Music, and the Brain: A Mysterious Relationship* (Cambridge: MIT Press) 141–172
- Krumhansl, Carol L. 1990. *Cognitive Foundations of Musical Pitch* (New York: Oxford University Press)
- Kusumoto, Kiyomi, and Elliott Moreton. 1997. 'Native Language Determines the Parsing of Nonlinguistic Rhythmic Stimuli', *The Journal of the Acoustical Society of America*, 102, 3204–3204
- Laden, Bernice. 1994. 'Melodic Anchoring and Tone Duration', *Music Perception: An Interdisciplinary Journal*, 12, 199–212
- Lai, Eric. 1999. 'Old Wine in New Bottles: The Use of Traditional Material in New Chinese Music', *ACMR Reports*, 1–22
- Langer, Susanne K. 1942. *Philosophy in A New Key* (Cambridge, MA: Harvard University Press)
- Langer, Susanne K. 1953. *Feeling and Form, A Theory of Art Developed from Philosophy in A New Key* (London: Routledge and Kegan Paul)
- Langus, Alan, Shima Seyed-Allaei, Ertuğrul Uysal, Sahar Pirmoradian, Caterina Marino, and Sina Asaadi, Ömer Eren, Juan M. Toro, Marcela Peña, Ricardo A. H. Bion, and Marina Nespore. 2016. 'Listening Natively Across Perceptual Domains?', *Journal of Experimental Psychology: Learning, Memory, And Cognition*, 42, 1127–1139

- Large, Edward W., Caroline Palmer, and Jordan B. Pollack. 1995. 'Reduced Memory Representations for Music', *Cognitive Science*, 19, 53–96
- Lee, Christopher S., and Neil P. Todd. 2004. 'Towards an Auditory Account of Speech Rhythm: Application of A Model of the Auditory 'Primal Sketch' To Two Multi-Language Corpora', *Cognition*, 93, 225–254
- Lee, Young-Min. 1988. 'The Development of Western-Style Orchestral Music in Korea' (PhD thesis, Catholic University of America)
- Leech-Wilkinson, Daniel. 2007. 'Sound and Meaning in Recordings of Schubert's 'Die Junge Nonne'', *Musicae Scientiae*, 11, 209–236
- Leech-Wilkinson, Daniel. 2009. *The Changing Sound of Music: Approaches to Studying Recorded Musical Performances* (London: CHARM)
<<http://www.charm.kcl.ac.uk/studies/chapters/chap1.html>> [Accessed 30 July 2019]
- Lehiste, Ilse. 1977. 'Isochrone Reconsidered', *Journal of Phonetics*, 5, 253–263
- Lerdahl, Fred, and Ray Jackendoff. 1983. *A Generative Theory of Tonal Music* (Cambridge, MA: MIT Press)
- Lerner, Murray. 1979. *From Mao To Mozart: Isaac Stern in China* (China: Hopewell Foundation)
- Lesznai, Lajos. 1961. *Bartók*, trans. by Percy M. Young (London: J. M. Dent & Sons)
- Levinson, Jerrold. 1997. *Music in the Moment* (Ithaca, NY: Cornell University Press)
- Li, Aike and Brechtje Post. 2014. 'L2 Acquisition of Prosodic Properties of Speech Rhythm: Evidence from L1 Mandarin and German Learners of English', *Studies in Second Language Acquisition*, 36, 223-255

- Li, Xiaole. 2003. 'Chen Yi's Piano Music: Chinese Aesthetics and Western Models' (PhD thesis, University of Hawaii at Manoa)
- Lieberman, Philip. 1967. *Intonation, Perception, And Language* (Cambridge, MA: MIT Press)
- List, George. 1963. 'The Boundaries of Speech and Song', *Ethnomusicology*, 7, 1–16
- Liu, Xiaoluan, and Yi Xu. 2015. 'Relations Between Affective Music and Speech: Evidence from Dynamics of Affective Piano Performance and Speech Production', *Frontiers in Psychology*, 6, 886
- Lomax, Alan. 1968. *Folk Song Style and Culture* (New Brunswick, NJ: Transaction Publisher)
- London, Justin, and Katherine Jones. 2011. 'Rhythmic Refinements to the nPVI Measure: A Reanalysis of Patel & Daniele (2003A)', *Music Perception: An Interdisciplinary Journal*, 29, 115–120
- London, Justin. 2001. 'Rhythm' in *Grove Music Online* (*Oxford Music Online*)
<<http://www.oxfordmusiconline.com/subscriber/article/grove/music/45963>>
[Accessed: 1 July 2017]
- Low, Ee Ling, Esther Grabe, and Francis Nolan. 2000. 'Quantitative Characterizations of Speech Rhythm: Syllable-Timing in Singapore English', *Language and Speech*, 43, 377–401
- Maema, Takanori, and Yūichi Iwano. 2001. *Nihon No Piano 100-Nen* (Tokyo: Soshisha)
- Marie, Céline, Cyrille Magne, and Mireille Besson. 2011. 'Musicians and the Metric Structure of Words', *Journal of Cognitive Neuroscience*, 23, 294–305
- Matravers, Derek. 1998. *Art and Emotion* (Oxford: Clarendon Press)
- McCole, John. 1993. *Walter Benjamin And the Antinomies of Tradition* (Ithaca, New York: Cornell University Press)

- McGowan, Rebecca W., and Andrea G. Levitt. 2011. 'A Comparison of Rhythm in English Dialects and Music', *Music Perception: An Interdisciplinary Journal*, 28, 307–314
- Mehl, Margaret. 2013. 'Western Art Music in Japan: A Success Story?', *Nineteenth-Century Music Review*, 10, 211–222
- Melvin, Sheila, and JinDong Cai. 2004. *Rhapsody in Red* (New York: Algora Publishing)
- Meyer, Leonard B. 1956. *Emotion and Meaning in Music* (Chicago: University of Chicago Press)
- Miller, Allan. 2000. *Musical Encounters* (China: Four Oaks Foundation)
- Milovanov, Riia, Minna Huotilainen, Paulo A.A. Esquef, Paavo Alku, Vesa Välimäki, and Mari Tervaniemi. 2009. 'The Role of Musical Aptitude and Language Skills in Preattentive Duration Processing in School-Aged Children', *Neuroscience Letters*, 460, 161–165
- Mithen, Steven J. 2005. *The Singing Neanderthals: The Origins of Music, Language, Mind and Body* (London: Weidenfeld & Nicholson)
- Mo, Yoonsook. 2008. 'Duration and Intensity as Perceptual Cues for Naïve Listeners' Prominence and Boundary Perception', in *Speech Prosody 2008 Conference* (Campinas, Brazil: International Speech Communication Association) 739–742 <https://www.isca-speech.org/archive/sp2008/papers/sp08_739.pdf> [Accessed 8 August 2019]
- Mok, Peggy P.K., and Volker Dellwo. 2008. 'Comparing Native and Non-Native Speech Rhythm Using Acoustic Rhythmic Measures: Cantonese, Beijing Mandarin and English', in *4th International Conference* (Campinas, Brazil: Speech Prosody 2008) 423–426
- Mor, Noam, Lior Wolf, Adam Polyak, and Yaniv Taigman. 2018. *A Universal Music Translation Network* (Computer Research Repository, CoRR) <<https://arxiv.org/abs/1805.07848>>

- Moran, Nikki. 2014. 'Social implications arise in embodied music cognition research which can counter musicological "individualism"', *Frontiers in Psychology*, 5
<<https://www.frontiersin.org/articles/10.3389/fpsyg.2014.00676/full>> [Accessed on 28 August 2019]
- Morley, David, and Kevin Robins. 1995. *Spaces of Identity: Global Media, Electronic Landscapes, and Cultural Boundaries* (London: Routledge) 147–173
- Nattiez, Jean-Jacques. 1973. 'Linguistics: A New Approach for Musical Analysis?', *International Review of the Aesthetics and Sociology of Music*, 4, 51
- Nazzi, Thierry, Josiane Bertoncini, and Jacques Mehler. 1998. 'Language Discrimination by Newborns: Toward an Understanding of the Role of Rhythm.', *Journal of Experimental Psychology: Human Perception and Performance*, 24, 756–766
- Nettl, Bruno. 1958. 'Some Linguistic Approaches to Musical Analysis', *Journal of the International Folk Music Council*, 10, 37
- Nolan, Francis, and Hae-Sung Jeon. 2014. 'Speech Rhythm: A Metaphor?', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369, 20130396
- Ohala, John J. 1984. 'An Ethological Perspective on Common Cross-Language Utilization of F0 of Voice', *Phonetica*, 41, 1–16
- Ohala, John J. 1994. 'The frequency codes underlie the sound symbolic use of voice pitch', in *Sound Symbolism*, ed. by Leanne Hinton, Johanna Nicolas, and John J. Ohala (Cambridge: Cambridge University Press) 325–347
- Ohala, John J. 1983. 'Cross-Language Use of Pitch: An Ethological View', *Phonetica*, 40, 1–18
- Ohgushi, K. 1999 'Physical Analysis of Piano Performances No. 2', *Spring Meeting of Acoustical Society of Japan* (Kanagawa, March) (in Japanese)

- Ohgushi, K. 2002. 'Comparison of Dotted Rhythm Expression between Japanese and Western Pianist', *7th International Conference on Music Perception and Cognition*, (Sydney)
- Oliveira, Jr., Miguel. 2003. 'Pitch Reset as A Cue for Narrative Segmentation', in *Proceedings of IP2003– Prosodic Interfaces* (Nantes, France: IP2003), 73–78
- Oram, Nicholas, and Lola L. Cuddy. 1995. 'Responsiveness of Western Adults to Pitch-Distributional Information in Melodic Sequences', *Psychological Research*, 57, 103–118
- Ott, Cyrill. 2011. 'Processing of Voiced and Unvoiced Acoustic Stimuli in Musicians', *Frontiers in Psychology*, 2, 195
- Palmer, Caroline, and Sean Hutchins. 2008. 'What Is Musical Prosody?', in *Psychology of Learning and Motivation: Advances in Research and Theory*, 46th edn., ed. by Brian Ross (San Diego, CA: Elsevier Academic Press) 245–278
- Palmer, Caroline, Melissa K. Blakeslee, and Peter W. Jusczyk. 1999. 'Recognition of Prosodic Cues in Music Performance', *The Journal of the Acoustical Society of America*, 106, 2235–2235
- Palmer, Caroline, Melissa K. Jungers, and Peter W. Jusczyk. 2001. 'Episodic Memory for Musical Prosody', *Journal of Memory and Language*, 45, 526–545
- Palmer, Caroline. 1989. 'Mapping Musical Thought to Musical Performance.', *Journal of Experimental Psychology: Human Perception and Performance*, 15, 331–346
- Palmer, Caroline. 1996. 'On the Assignment of Structure in Music Performance', *Music Perception: An Interdisciplinary Journal*, 14, 23–56
- Patel, Aniruddh D, and Joseph R Daniele. 2003. 'An Empirical Comparison of Rhythm in Language and Music', *Cognition*, 87, B35–B45

- Patel, Aniruddh D. 2008. *Music, Language, and the Brain* (New York: Oxford University Press)
- Patel, Aniruddh D. 2011. 'Language, Music, and the Brain: A Resource-Sharing Framework', in *Language and Music as Cognitive Systems* (Oxford: Oxford University Press) 204–223
- Patel, Aniruddh D. 2011. 'Why Would Musical Training Benefit the Neural Encoding of Speech? The OPERA Hypothesis', *Frontiers in Psychology*, 2, 142
- Patel, Aniruddh D., and Emily Morgan. 2016. 'Exploring Cognitive Relations Between Prediction in Language and Music', *Cognitive Science*, 41, 303–320
- Patel, Aniruddh D., John R. Iversen, and Jason C. Rosenberg. 2006. 'Comparing the Rhythm and Melody of Speech and Music: The Case of British English And French', *Journal of the Acoustical Society of America*, 119, 3034–3047
- Pearl, Jonathan G. Secora. 2006. 'Eavesdropping with a Master: Leoš Janáček and the Music of Speech', *Empirical Musicology Review*, 1, 131–165
- Pell, Marc D. 2001. 'Influence of Emotion and Focus Location on Prosody in Matched Statements and Questions', *The Journal of the Acoustical Society of America*, 109, 1668–1680
- Peretz, Isabelle, and Max Coltheart. 2003. 'Modularity of Music Processing', *Nature Neuroscience*, 6, 688–691
- Peretz, Isabelle. 2009. 'Music, Language and Modularity Framed in Action', *Psychologica Belgica*, 49, 157
- Petkova, Diana. 2005. 'Cultural Identity in a Pluralistic World', in *Cultural Identity in an Intercultural Context*, ed. by Diana Petkova and Jaako Lehtonen (University of Jyväskylä) 11–66

- Peynircioğlu, Zehra F., Wenyan Bi, and William Brent. 2018. 'The 'Asian Bias' Illusion in Musical Performance: Influence of Visual Information', *American Journal of Psychology*, 131, 295–305
- Pierrehumbert, Janet B. 1999. 'Prosody and Intonation', in *MIT Encyclopedia of the cognitive sciences*, ed. by Robert A. Wilson and Frank C. Keil, (Cambridge, MA: MIT Press) 679–682
- Pike, Kenneth L. 1945. *The Intonation of American English* (Ann Arbor: University of Michigan Press)
- Pisoni, David B. 1997. 'Some thoughts on 'normalization' in speech perception', in *Talker Variability in Speech Processing*, ed. by Keith Johnson and John W. Mullennix (San Diego, CA: Academic Press) 9–32
- Plato. 2000. *The Republic*, ed. by G. R. F. Ferrari, trans. by Tom Griffith (Cambridge: Cambridge University Press)
- Post, Brechtje, and Elinor Payne. 2018. 'Speech Rhythm in Development', in *The Development of Prosody in First Language Acquisition*, ed. by Pilar Prieto and Nürria Esteve-Gibert (John Benjamins Publishing Company) 125–144 <<https://www.jbe-platform.com/content/books/9789027264213-tilar.23.07pos>> [Accessed 29 July 2019]
- Prockup, Matthew, David Grunberg, Alex Hrybyk, and Youngmoo E. Kim. 2013. 'Orchestral Performance Companion: Using Real-Time Audio to Score Alignment', *IEEE Multimedia*, 20, 52–60
- Raffman, Diana. 1993. *Language, Music and Mind* (Cambridge, MA: MIT Press)
- Raju, Marju, Eva Liina Asu, and Jaan Ross. 2010. 'Comparison of Rhythm in Musical Scores and Performances as Measured with the Pairwise Variability Index', *Musicae Scientiae*, 14, 51–71

- Rammsayer, Thomas, and Eckart Altenmüller. 2006. 'Temporal Information Processing in Musicians and Nonmusicians', *Music Perception*, 24, 37–48
- Ramus, Franck, Marina Nespore, and Jacques Mehler. 1999. 'Correlates of Linguistic Rhythm in the Speech Signal', *Cognition*, 73, 265–292
- Ramus, Frank. 2002. 'Acoustic Correlates of Linguistic Rhythm: Perspectives', in *Proceedings of Speech Prosody 2002, Aix-En-Provence* (Aix-en-Provence: Laboratoire Parole et Langage) 115–120
- Rebuschat, Patrick, Martin Rohmeier, John A. Hawkins, and Ian Cross. 2011. *Language and Music as Cognitive Systems* (Oxford: Oxford University Press)
- Repp, Bruno H. 1992. 'Diversity and Commonality in Music Performance: An Analysis of Timing Microstructure in Schumann's "Träumerei"', *The Journal of the Acoustical Society of America*, 92, 2546–2568
- Repp, Bruno H. 1995. 'Quantitative Effects of Global Tempo on Expressive Timing in Music Performance: Some Perceptual Evidence', *Music Perception: An Interdisciplinary Journal*, 13, 39–57
- Repp, Bruno H. 2000. 'Pattern Typicality and Dimensional Interactions in Pianists' Imitation of Expressive Timing and Dynamics', *Music Perception: An Interdisciplinary Journal*, 18, 173–211
- ResearchInChina. 2014) Global and China Piano Industry Report, 2014-2017. Available from: <<http://www.researchinchina.com/Htmls/Report/2014/7964.html>> [Accessed: 21 April 2017]
- ResearchMoz. 2013. 'Global and China Piano Industry 2013: Industry Analysis, Size, Share, Growth, Trends and Forecast Research Report' <<http://www.prweb.com/releases/2013/7/prweb10945813.htm>> [Accessed: 21 April 2017]

- Ridley, Aaron. 1995. *Music, Value, and the Passions* (Ithaca: Cornell University Press)
- Ritterman, Janet. 2004. *Knowing more than we can tell: Artistic practice and integrity*.
Queensland Conservatorium Griffith University.
- Roach, Peter. 1982. 'On the distinction between 'stress-timed' and 'syllable-timed' languages', in *Linguistic Controversies*, ed. by David Crystal (London: Edward Arnold) 73–79
- Roth, Henry. 1997. *Violin Virtuosos* (Los Angeles, CA: California Classics Books)
- Rousseau, Jean-Jacques. 2009. *Essay on the Origin of Languages and Writings Related to Music*, trans. by John T. Scott (Dartmouth, NH: Dartmouth College Press)
- Rubinstein, Anton. 1892. *Music and Its Masters: A Conversation on Music* (London: Augener) <<https://archive.org/details/musicitsmastersc00rubiuoft/page/n8>> [Accessed 28 July 2019]
- Ruwet, Nicholas. 1967. 'Linguistics and Musicology', *International Social Science Journal*, 19, 79–87
- Sadakata, Makiko, and Kaoru Sekiyama. 2011. 'Enhanced Perception of Various Linguistic Features by Musicians: A Cross-Linguistic Study', *Acta Psychologica*, 138, 1–10
- Schack, Katrina. 2000. 'Comparison of intonation patterns in Mandarin and English for a particular speaker', in *University of Rochester Working Papers in the Language Sciences, Vol. Spring, no. 1*, ed. by Katherine M. Crosswhite and Joyce McDonough, 24–55
- Schiavio, Andrea, Dylan van der Schyff, Michele Biasutti, Nikki Moran, and Richard Parncutt. 2019. 'Instrumental Technique, Expressivity, and Communication. A Qualitative Study on Learning Music in Individual and Collective Settings', *Frontiers in Psychology*, 10

<<https://www.frontiersin.org/articles/10.3389/fpsyg.2019.00737/full#B3>> [Accessed 28 August 2019]

Schneider, Christopher J. 2010. 'Introduction: Music and Identity', *Studies in Symbolic Interaction*, 35, 7–10

Schopenhauer, Arthur. 1958. 'Die Welt als Wille und Vorstellung, The World as Will and Representation', trans. by Eric F. J. Payne (Indian Hills, CO: Falcon's Wing Press)

Schreuder, Maartje. 2006. 'Prosodic Processes in Language and Music' (PhD thesis, University of Groningen)

Shaffer, L. Henry and Neil P. M. Todd. 1987. 'The interpretive component in musical performance', in *Action and Perception in Rhythm and Music*, ed. by Alf Gabrielsson (Stockholm: Royal Swedish Academy of Music) 139–152

Sherzer, Joel, and Anthony C Woodbury. 1987. *Native American Discourse: Poetics and Rhetoric* (Cambridge: Cambridge University Press)

Slevc, L. Robert, and Akira Miyake. 2006. 'Individual Differences in Second-Language Proficiency', *Psychological Science*, 17, 675–681

Slevc, L. Robert, and Aniruddh D. Patel. 2011. 'Meaning in Music and Language: Three Key Differences', *Physics of Life Reviews*, 8, 110–1, 125–8

Sloboda, J. 2004. *Exploring the Musical Mind: Cognition, Emotion, Ability, Function*. (New York: Oxford University Press)

Sloboda, John A. 1976. 'The Effect of Item Position on the Likelihood of Identification by Inference in Prose Reading and Music Reading.', *Canadian Journal of Psychology*, 30, 228–237

Sloboda, John A. 1983. 'The Communication of Musical Metre In Piano Performance', *Quarterly Journal of Experimental Psychology Section A*, 35, 377–396

- Sloboda, John A. 1986. *The Musical Mind: The Cognitive Psychology of Music* (New York: Oxford University Press)
- Slobodian, Lydia N. 2008. 'Perception and Production of Linguistic and Musical Rhythm by Korean And English Middle School Students', *Empirical Musicology Review*, 3, 187–204
- Springer, George P. 1956. 'Language and Music: Parallels and Divergences', in *For Roman Jakobson* (The Hague) 504–513
- Stern, Theodore. 1957. 'Drum and Whistle 'Languages': An Analysis of Speech Surrogates', *American Anthropologist*, 59, 487–506
- Strait, Dana L., and Nina Kraus. 2011. 'Can You Hear Me Now? Musical Training Shapes Functional Brain Networks for Selective Auditory Attention and Hearing Speech in Noise', *Frontiers in Psychology*, 2, 113
- Stravinsky, Igor. 1947. *Poetics of Music: In the Form of Six Lessons*, trans. by Arthur Knodel and Ingolf Dahl (Cambridge, MA: Harvard University Press)
- Stravinsky, Igor. 1998. *An Autobiography* (New York: W. W. Norton & Company)
- Suchoff, Benjamin (ed.). 1976. 'Harvard Lectures' in *Béla Bartók Essays* (New York: St. Martin's Press) 377–381
- Sundberg, Johan, and Björn Lindblom. 1976. 'Generative Theories in Language and Music Descriptions', *Cognition*, 4, 99–122
- Sundberg, Johan, and Lars Frydén. 1985. 'Teaching A Computer to Play Melodies Musically', *Analytica: Festschrifts for Ingmar Bengtsson*, 47, 67–76
- Sundberg, Johan, and Violet Verrillo. 1980. 'On the Anatomy of the Ritard: A Study of Timing in Music', *The Journal of the Acoustical Society of America*, 68, 772–779

- Sundberg, Johan, Anders Askenfelt, and Lars Fryden. 1983. 'Musical Performance: A Synthesis-By-Rule Approach', *Computer Music Journal*, 7, 37
- Sundberg, Johan, Anders Friberg, and Lars Fryden. 1991. 'Threshold and Preference Quantities of Rules for Music Performance', *Music Perception: An Interdisciplinary Journal*, 9, 71–91
- Sundberg, Johan. 1989. 'Aspects of Music Communication as Revealed by Analysis by Synthesis of Musical Performance', in *1st International Conference on Music Perception and Cognition* (Kyoto) 5–10
- Swain, Joseph P. 1997. *Musical Languages* (New York: Norton)
- Tânia, Lisboa, Aaron Williamon, Massimo Zicari, and Hubert Eiholzer. 2005. 'Mastery Through Imitation: A Preliminary Study', *Musicae Scientiae*, 9, 75–110
- Taruskin, Richard. 1988. 'The Pastness Of the Present and the Presence of the Past', in *Authenticity and Early Music: A Symposium*, ed. by Nicholas Kenyan (Oxford; New York: Oxford University Press) 137–210
- Tedlock, Dennis. 1983. *The Spoken Word and the Work of Interpretation* (Philadelphia: University of Pennsylvania Press)
- Temperley, Nicholas, and David Temperley. 2011. 'Music-Language Correlations and the 'Scotch Snap'', *Music Perception: An Interdisciplinary Journal*, 29, 51–63
- Thiemel, Matthias. 2001. 'Accent' in *Grove Music Online* (Oxford Music Online)
<<https://www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000000098>> [Accessed 9 August 2019]
- Tillmann, Barbara, Denis Burnham, Sebastien Nguyen, Nicolas Grimault, Nathalie Gosselin, and Isabelle Peretz. 2011. 'Congenital Amusia (Or Tone-Deafness) Interferes with Pitch Processing in Tone Languages', *Frontiers in Psychology*, 2, 120

- Timmers, Renee, and Makiko Sadakata. 2014. 'Training Expressive Performance by Means of Visual Feedback: Existing and Potential Applications of Performance Measurement Techniques', in *Expressiveness in Musical Performance*, ed. by Dorottya Fabian, Renee Timmers, and Emery Schubert (Oxford: Oxford University Press) 304–330
- Todd, Neil P. M. 1985. 'A Model of Expressive Timing in Tonal Music', *Music Perception: An Interdisciplinary Journal*, 3, 33–57
- Todd, Neil P. M. 1992. 'The Dynamics of Dynamics: A Model of Musical Expression', *Journal of the Acoustical Society of America*, 91, 3540–3550
- Todd, Neil P. M. 1995. 'The Kinematics of Musical Expression', *Journal of the Acoustical Society of America*, 97, 1940–1949
- Tokita, A. 2010. 'The Piano and Cultural Modernity in East Asia'. In *Philosophical and Cultural Theories of Music*, ed. by Eduardo de la Fuente and Peter Murphy, 221–242.
- Türk, Daniel G. 1982. *School of Clavier Playing or Instructions in Playing the Clavier for Teachers and Students*, trans. by Raymond H. Hagg (Lincoln: University of Nebraska Press)
- Umiker, Donna J. 1974. 'Speech Surrogates: Drum and Whistle Systems', in *Current Trends in Linguistics, 12: Linguistics and Adjacent Arts and Sciences* (The Hague: Mouton) 497–536
- Wade, Bonnie C. 2005. *Music in Japan: Experiencing Music, Expressing Culture* (New York: Oxford University Press)
- Wagner, Roy. 1981. *The Invention of Culture* (Chicago: University of Chicago Press)
- Wall, Sarah. 2006. 'An Autoethnography On Learning About Autoethnography', *International Journal of Qualitative Methods*, 5, 146–160

- Wall, Sarah. 2016. 'Toward A Moderate Autoethnography', *International Journal of Qualitative Methods*, 15
 <<https://journals.sagepub.com/doi/full/10.1177/1609406916674966#articleCitationDownloadContainer>> [Accessed on 23 August 2019]
- Walsh, Michael. 1983. 'Like A Flower on A Pond': The Classics Flourish in Japan, But How Deep Are Their Roots?', *Time Magazine*
- Wang, Grace. 2015. *Soundtracks of Asian America: Navigating Race Through Musical Performance* (Durham, NC: Duke University Press)
- Weber, William. 1975. *Music and the Middle Class: The Social Structure of Concert Life in London, Paris, and Vienna* (New York: Holmes & Meier)
- Weber, William. 2008. *The Great Transformation of Musical Taste: Concert Programming from Haydn to Brahms* (New York: Cambridge University Press)
- Wenk, Brian J. 1982. 'Speech Patterns in Music: the French (Rhythmic) Connection', *International Review of the Aesthetics and Sociology of Music*, 13, 191
- Wenk, Brian J. 1987. 'Just in Time: On Speech Rhythms in Music', *Linguistics*, 25
- White, Laurence, and Sven L. Mattys. 2007. 'Calibrating Rhythm: First Language and Second Language Studies', *Journal of Phonetics*, 35, 501–522
- White, Laurence, Elinor Payne, and Sven L. Mattys. 2009. 'Rhythmic and prosodic contrast in Venetan and Sicilian Italian', in *Phonetics and Phonology: Interactions and Interrelations*, ed. by Marina Vigário, Sónia Frota, and M. João Freitas (Amsterdam: John Benjamins) 137–158
- White, Laurence, Sven L. Mattys, and Lukas Wiget. 2012. 'Language Categorization by Adults Is Based on Sensitivity to Durational Cues, Not Rhythm Class', *Journal of Memory and Language*, 66, 665–679

- White, Laurence. 2014. 'Communicative Function and Prosodic Form in Speech Timing', *Speech Communication*, 63–64, 38–54
- Wiget, Lukas, Laurence White, Barbara Schuppler, Izabelle Grenon, Olesya Rauch, and Sven L. Mattys. 2010. 'How Stable Are Acoustic Metrics of Contrastive Speech Rhythm?', *Journal of the Acoustical Society of America*, 127, 1559–1569
- Woodbury, Anthony C. 1987. 'Rhetorical Structure in A Central Alaskan Yupik Eskimo Traditional Narrative', in *Native American Discourse* (Cambridge: Cambridge University Press) 176–239
- Woodbury, Anthony C. 1992. 'Prosodic Elements and Prosodic Structures in Natural Discourse', in *Proceedings of the IRCS Workshop on Prosody in Natural Speech, IRCS Report 92* (University of Pennsylvania, Institute for Research in Cognitive Science) 241–253
- Woodrow, Herbert. 1909. *A Quantitative Study of Rhythm: The Effect of Variations in Intensity, Rate and Duration* (New York: Science Press)
- Wray, Alison. 1998. 'Protolanguage as A Holistic System for Social Interaction', *Language & Communication*, 18, 47–67
- Wray, Alison. 2000. 'Holistic Utterances in Protolanguage: The Link from Primates to Humans', in *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form* (New York: Cambridge University Press) 285–302
- Xu, Mo. 2018. 'The High Finger Piano Technique in China: past, present, and future' (PhD thesis, University of Iowa)
- Yang, Mina. 2014. *Planet Beethoven: Classical Music at The Turn of the Millennium* (Middletown, CT: Wesleyan University Press)

- Yee, William, Susan Holleran, and Mari Riess Jones. 1994. 'Sensitivity to Event Timing in Regular and Irregular Sequences: Influences of Musical Skill', *Perception & Psychophysics*, 56, 461–471
- Yoshihara, Mari. 2007. *Musicians from A Different Shore: Asians And Asian Americans In Classical Music* (Philadelphia, PA: Temple University Press)
- Yung, Bell. 1983. 'Creative Process in Cantonese Opera II: The Process of T'ien Tz'u (Text-Setting)', *Ethnomusicology*, 27, 297–318
- Zhang, Xinting. 2012. 'A Comparison of Cue-Weighting in the Perception of Prosodic Phrase Boundaries in English And Chinese' (PhD thesis, University of Michigan)
- Zhang, Yanhong, Shawn L. Nissen, and Alexander L. Francis. 2008. 'Acoustic Characteristics of English Lexical Stress Produced by Native Mandarin Speakers', *Journal of the Acoustical Society of America*, 123, 4498–4513
- Zhao, Yilu. 2002. 'Cultural Divide Over Parental Discipline', *The New York Times* <<https://www.nytimes.com/2002/05/29/nyregion/cultural-divide-over-parental-discipline.html>> [Accessed 28 July 2019]
- Zimmermann, Elke, Lisette M. C. Leliveld, and Simone Schehka. 2013. 'Toward the evolutionary roots of affective prosody in human acoustic communication: a comparative approach to mammalian voices', in *Evolution of emotional communication: from sounds in nonhuman mammals to speech and music in man*, ed. by Eckart Altenmüller, Sabine Schmidt, and Elke Zimmermann (Oxford: Oxford University Press) 116–132